ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc. THESIS

Mahmood Najeeb ABED

SLEEP APNEA EVENTS DETECTION FROM POLYSOMNOGRAM STUDIES USING DEEP LEARNING TECHNIQUES

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

ADANA-2020

ABSTRACT

MSc. THESIS

SLEEP APNEA EVENTS DETECTION FROM POLYSOMNOGRAM STUDIES USING DEEP LEARNING TECHNIQUES

Mahmood Najeeb ABED

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

Supervisor : Assoc. Prof. Dr. Turgay İBRİKÇİ
Year: 2020, Pages: 71
Jury : Assoc. Prof. Dr. Turgay İBRİKÇİ
: Assoc. Prof. Dr. Sami ARICA
: Assoc. Prof. Dr. İrem Ersöz KAYA

The significant raising of deep learning techniques has made us discover its capability in the detection of Sleep Apnea (SA). Sleep Apnea is a respiratory issue described by irregular breathing samples during sleep, including loud snoring and successive waking. Uncountable sleep apnea cases are under-evaluated because of the remoteness, burden, or cost of Polysomnogram in sleep laboratories. Thus, automated detection has been applied using different machine learning methods. In this thesis, we proved the clinical presumption by targeting specific physiological signals, which are Chest, Abdominal, Oxygen Saturation (SpO_2). We were able to detect sleep apnea events from those signals and normal events as well. Two databases from Physionet are used in this thesis. The first database was a competition in 2018, called You Snooze You Win (YSYW), and the other database was also a competition in 2000, called the Apnea-ECG. We used the first one for training and testing, and for generalization, we used the second database for testing also. Epochs of 60 seconds in size have been handled as input datasets. Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) were included to be exploited in this task. Our results showed that the hybrid model consisting of Convolution Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) outperformed the other deep learning techniques. The top overall accuracy reached on the first database is 97%, and the second database is 92%. Finally, we discussed the obtained results with other related researches.

Keywords: Sleep Apnea, Polysomnogram, Deep Learning

ÖZ

YÜKSEK LİSANS TEZİ

UYKU APNE ETKİNLİKLERİ DERİN ÖĞRENME TEKNİKLERİ KULLANILAN POLİSOMNOGRAM ARAŞTIRMALARINDAN ALGILAMA

Mahmood Najeeb ABED

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK - ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Doç. Prof. Dr. Turgay İBRİKÇİ
Yıl: 2020, Sayfa: 71
Jüri : Doç. Prof. Dr. Turgay İBRİKÇİ
: Doç. Dr. Sami ARICA
: Doç. Dr. İrem Ersöz KAYA

Uyku apnesinin tespit etme kabiliyetini keşfetmemizi sağlayan derin öğrenme tekniklerinin önemli ölçümde artmasından kaynaklıdır. Uyku apnesi yüksek sesle horlama ve ardışık uyanma dahil olarak uyku sırasında düzensiz solunum örnekleri ile tanımlanan bir solunum sorunudur. Bir-çok sayıda değerlendirilmeyen uyku apnesinin nedeni plisoygramın uyku laboratuvarlarında uzak ve yüksek maliyete uygulanmasından kaynaklıdır. Bu tezde, belirli fizyolojik sinyalleri hedefleyerek klinik varsayımını kanıtladık, belirli olan fizyolojik sinyaller, Göğüs, Karın, Oksijen Doygunluğu (SpO2)dan algılanmaktadır, Bu sinyallerden ve normal olaylardan uyku apnesi olaylarını da tespit edebildik. Bu tezde iki veri tabanı Physionet kaynaklarını ele alarak kullanılmıştır. Ilk veri tabanı 2018'de You Snooze You Win (YSYW) adlı, ve ikinci veri tabanı da 2000 yılında Apne-EKG adlı. Derin öğrenme tekniklerinin önemli ölçüde artması, Uyku Apnesinin (SA) saptanmasındaki kabiliyetini tespit etmesini sağladı. Uyku apnesi bir solunum sorunu olarak görünmektedir, ilk verileri eğtim ile test için kullanılırken ikinci veri tabanı ise test için kullanıldı. 60 saniyelik dönemler, giriş veri kümeleri olarak işlendi. Geçitli Tekrarlayan Birim (GRU) ve Uzun Kısa Süreli Bellek (LSTM) bu göreve faydalanmak için dahil edildi. Çalışma Sonuçlarımız, Konvolüsyon Sinir Ağları (CNN'ler) ve Tekrarlayan Sinir Ağlarından (RNN'ler) oluşan hibrit modelin diğer derin öğrenme tekniklerinden daha iyi performans gösterdiğini gösterdi. İlk veritabanındaki en yüksek genel doğruluk% 97' iken, ikinci veritabanındaki genel doğruluk oranını % 92 elde edildi. Sonuç olarak elde ediler veriler değer araştırma sonucları ile karşılaştırılıp tartışıldı.

Anahtar Kelimeler: Uyku Apnesi, Polisomnogram, Derin Öğrenme

EXTENDED SUMMARY

Currently, while the world is going through with a global health crisis related to the Coronavirus, Artificial Intelligence (AI) has been proven that it is one of the best available resources to control the disease. AI techniques indicate that the lives of people will dramatically change in the next decade. People will start to depend on automation methods in many fields, especially in healthcare.

Researchers and doctors can discover patterns in the collected data and linked them up from many patients. That helps them to figure out different behaviors of predicting or diagnosing illnesses and find conducts to progress clinical care. Then the information can be used to understand more about disease risks and causes, improving diagnosis, developing new treatments, and preventing the disease.

Incessantly, people deal with millions of data in either numerical, categorical, or image forms. That data includes essential information through which a specific model can be obtained to understand what is happening, what happened, or even what will happen to a particular case. The biomedical data is very valuable as the models derived from it will invest in providing specific treatments, minimizing the cost of the medical diagnosis, or even diagnosing more accurately.

Sleep apnea is a respiratory-related disease, and its diagnosis requires effort because it occurs only during the patient's sleep. It includes three major types, and the most critical one is Obstructive Sleep Apnea (OSA). The patient is required to go to the sleep lab and make a sleep study that includes sleeping for an entire night in the laboratory for the diagnosis. Perhaps it will be an uncomfortable night due to the connection of electrodes with different positions of the body to obtain the necessary physiological signals. Sometimes, it requires that the patient has to come back again to sleep one more night for treatment, as the currently available treatment is Continuous Positive Airway Pressure (CPAP) titration, which controls the airway and keeps it open continuously.

We used two types of sleep apnea databases. These databases are labeled for various diseases related to breathing and sleep. We performed data preprocessing and cleaning, where we chose to deal with the binary classification, and this means that all other windows for apnea types were characterized as just apnea. Moreover, we selected only three signals from every database for two reasons. The first is the ease of obtaining these signals through some devices commercially available. The second reason is that by observing these signals, we get apparent signs indicating the presence or absence of sleep apnea.

Deep learning is a subset of neural networks that are the foundation of AI. It has effective techniques in multiple tasks, mainly supervised learning. Of course, we can’t cover all deep learning techniques, as it is constantly evolving and significant. There are thousands of research papers that are published continuously on new methods and architectures. Therefore, we have driven to cover the two main significant types, which are the Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). We have also made simple and effective merging to obtain a hybrid model that is used in carrying out the process of detecting sleep apnea events with a promised accuracy.

Likewise, we used another database, but we did not train a model on it. Instead, we used the previous model from each technique for testing on it. The project has also been implemented and executed entirely on the Google Colab platform, which is built for the development of Python codes. The reason is due to the presence of a high GPU card and a large RAM in addition to the super internet speed that enabled us to download databases of 135 GB within one hour or less. We evaluated all models according to the well-known metrics extracted from the confusion matrix, like accuracy, precision, recall, and f1-score.

By reviewing the previous related researches in using deep learning, particularly, we concluded that our results are very acceptable concerning the databases used, the number of signals used, and the trained model approach.

Finally, Automated technology has redefined healthcare and medical aid in the past decade with additional importance given to big data analytics and user-friendly software applications. Hence, sleep apnea detection will be the first real step in this field for us, but not the last.

GENİŞLETİLMİŞ ÖZET

Şu anda, dünya Coronavirüs ile ilgili küresel bir sağlık krizi yaşarken, Yapay Zeka'nın (AI) hastalığı kontrol etmek için mevcut olan en iyi kaynaklardan biri olduğu kanıtlandı.Yapay zeka teknikleri, insanların hayatlarının önümüzdeki on yılda dramatik olarak değişeceğini gösteriyor.İnsanlar birçok alanda otomasyon yöntemlerine güvenmeye başlayacak.özellikle sağlık hizmetlerinde.

Araştırmacılar ve doktorlar, toplanan verilerdeki desenleri keşfedebilir ve bunları birçok hastaya bağlayıp. Bu, hastalıkları öngörmek veya teşhis etmek için farklı davranışları anlamalarına ve klinik bakımı ilerletmek için yöntemler bulmalarına yardımcı olacaktır. Daha sonra bu bilgiler, hastalık riskleri ve nedenleri hakkında daha fazla bilgi edinmek, tanıyı iyileştirmek, yeni tedaviler geliştirmek ve hastalığı önlemek için kullanılabilir olacaktır.

İnsanlar sürekli olarak sayısal, kategorik veya görüntü formları seklinde olan milyonlarca verilereler ile karşılaşıp uğraşıyor ve ilgileniyor. Bu veriler, neler olduğunu, ne olduğunu ve hatta belirli bir vakaya ne olacağını anlamak için belirli bir modelin elde edilebileceği temel bilgileri içermektedir. Biyomedikal veriler özel tedavi modülleri olarak ve tıbbi teşhisin maliyetini en aza indirmek için Ya da daha doğru teşhis koyabilmek için çok değerli bir veridir.

Uyku apnesi solunumla ilgili bir hastalıktır ve teşhis konulması zor olan bir hastaliktir çünkü sadece hastanın uyku sırasında ortaya çıkar .uyku apnesi üç ana tip içerir ve en kritik olanı Obstrüktif Uyku Apnesi'dir (OSA). Hastanı teşhis edebilmek için uyku laboratuvarına gitmeyi ve laboratuvarda bütün gece uyumayi içeren bir uyku çalışma yapması gerekmektedir. Gerekli fizyolojik sinyalleri alabilmek için elektrotların vücudun farklı pozisyonlarına bağlanması gerekmektedir bu nedenle hasta için rahatsız edici bir gece olabilir. Bazan hastanın tedavi alabilmesi için bir geceden fazla uyuması gerekebilir. Şu anda mevcut tedavi, hava yolunu kontrol eden ve sürekli açık tutan Pozitif Hava Yolu Basıncı titrasyonudur (CPAP).

İki tür uyku apnesi veri tabanı kullandık. Bu veritabanları, solunum ve uyku ile ilgili çeşitli hastalıklar için etiketlenmiştir.Veri ön işleme ve temizleme işlemini gerçekleştirdik ikili sınıflandırma ile ilgilenmeyi seçtik ve bu apne tipleri için diğer tüm pencerelerin sadece apne olarak karakterize edildiği anlamına gelir iki nedenden ötürü her veritabanından yalnızca üç sinyal seçtik. Birinci neden bu sinyalleri ticari olarak temin edilebilen bazı cihazlarla elde etmenin kolaylığı, ikincisi ise bu sinyalleri gözlemleyerek uyku apnesinin varlığını veya yokluğunu gösteren belirgin işaretler gözlemleyerek kanitlanması.

Derin öğrenme, yapay zekanın temelini oluşturan sinir ağlarının bir alt kümesidir. Birden çok görev ve etkili tekniklere sahiptir, özelikle denetimli öğrenme. Yapay Zekanın Sürekli gelişmekte olması için ve önemli olduğundan dolayı, tüm derin öğrenme yöntemlerini ele alıp üzerine konuşmamız mümkün değildir. Sürekli olarak binlerce araştırma makalesi yeni yöntemler ve mimariler ile ilgili yayınlamaktadır. Bu nedenle, Evrişimli Sinir Ağları (CNN'ler) ve Tekrarlayan Sinir Ağları (RNN'ler) olan iki ana olan önemli türkülerini kapsamaya çalıştık. Ayrıca uyku apnesi olaylarını vaat edilen bir doğrulukla tespit etme sürecini gerçekleştirmede kullanılan hibrit bir model elde etmek için basit ve etkili birleştirme yapmaya çalıştık.

Aynı şekilde başka bir veritabanı kullandık, ancak bunun üzerine bir model eğitmedik. Bunun yerine, test etmek için her teknikten önceki modeli kullanıp sonuçlarını elde etmeye çalıştık. Proje tamamen Google Colab platformunda uygulanıp gerçekleştirildi. Google Colab platformu Python kodlarının geliştirilmesi için oluşturulmuştur. Bunun nedeni, 135 GB veri tabanlarını bir saat veya daha kısa sürede indirmemizi sağlayan süper internet hızında ek olarak yüksek bir GPU kartının ve büyük bir RAM'in varlığından kaynaklanıyor. Tüm modelleri doğruluk, kesinlik, geri çağırma ve F1 skoru gibi kafa karışıklığı matrisinden çıkarılan iyi bilinen metriklere göre değerlendirdik.

Özelikle önceki derin öğrenme araştırmaları inceleyerek sonuçlarımızın kullanılan veritabanları ve kullanılan sinyal sayısı ve eğitimli model yaklaşımı açısından çok kabul edilebilir olduğu sonucuna vardık.

Son olarak, otomatikleştirilmiş teknoloji, büyük veri analitiğine ve kullanıcı dostu yazılım uygulamalarına verilen ek önem ile son on yılda sağlık ve tıbbi yardımı yeniden tanımlamıştır. Dolayısıyla uyku apnesinin tespiti bizim için bu alandaki ilk gerçek adım olacak ama son değil.

ACKNOWLEDGMENTS

I want to express my deepest gratitude to Assoc. Prof. Dr. Turgay İBRİKÇİ for his supervision, direction, motivation, constructive advice, and continued faith in me in this thesis. It was a huge honor for me to have Assoc. Prof. Dr. Turgay İBRİKÇİ as a supervisor.

I also extend my gratitude to Assoc. Prof. Dr. Sami ARICA for his support and motivation during the academic year and for his approval to be a member of the jury.

I want to thank Assoc. Prof. Dr. İrem Ersöz KAYA, as well for her consent to be a member of the jury.

I owe extra appreciation to Assoc. Prof. Dr. Sedat KULECİ for sharing his knowledge and experience on sleep apnea.

I would also like to thank Mr. Larry Brewer, the owner of the first sleep academy, for supporting me a great deal by giving me free access to his educational website tutorials.

I want to share my sincere thanks and gratitude to the Turkish Government and people for giving me this opportunity and for their hospitality.

I would like to express big gratitude to my diamond planet for having her always beside me on this challenging journey away from home and our families.

Finally, I wish to express my heartfelt thankfulness to my family for their patience, encouragement, and continuous moral support.

CONTENTS **PAGE**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

LIST OF ABBREVIATIONS

ASV	:	Adaptive Servo Ventilation
AlexNet	:	Alex Krizhevsky Network
AASM	:	American Academy of Sleep Medicine
AUROC	:	Area Under Receiver Operating Characteristic
AI	:	Artificial Intelligence
BiPAP	:	Bilevel Positive Airway Pressure
CSA	:	Central Sleep Apnea
CDAC	:	Clinical Data Animation Center
COLAB	:	Collaboratory
CCNL	:	Computational Clinical Neurophysiology Laboratory
CuDNN	:	Compute Unified Device Architecture Deep Neural Network
		Continuous Positive Airway Pressure
CPAP	:	Convolutional Neural Network
CNN	:	Deep Belief Network
DBN	:	Deep Learning
DL	:	Deep Neural Network
DNN	:	Electrocardiography
ECG	:	Electroencephalography
EEG	:	Electromyography
EMG	:	Electrooculography
EOG	:	European Data Format
EDF	:	False Negative
FN	:	False Positive
FP	:	Gated Recurrent Unit
GRU	:	Graphics Processing Unit
GPU	:	Heart Rate Variability

HRV	:	Hidden Markov Model
HMM	:	Instantaneous Heart Rate
IHR	:	Long Short-Term Memory
LSTM	:	Massachusetts General Hospital
MGH	:	Machine Learning
ML	:	Massachusetts Institute of Technology-Beth Israel Hospital
MIT-BIH	:	Multi-Ethnic Study of Atherosclerosis Obstructive Sleep Apnea
		Osteoporotic Fractures in Men
MESA	:	Oxygen Saturation
OSA	:	Quadratic Discriminant
MrOS	:	Polysomnography
SpO2	:	Random-Access Memory
QA	:	Rapid Eye Movement
PSG	:	Rectified Linear Unit
RAM	:	Recurrent Neural Network
REM	:	Residual Neural Network
ReLu	:	Respiratory Effort-Related Arousals
RNN	:	Seoul National University Hospital
ResNet	:	Sleep Heart Health Study
RERA	:	Sleep Apnea
SNUH	:	Support Vector Machine
SHHS	:	True Negative
SA	:	True Positive
SVM	:	University College Dublin
TN	:	Visual Geometry Group
TP	:	Waveform-Database
UCD	:	Yann LeCun Network
VGG	:	You Snooze You Win

WFDB :

LeNet :

YSYW :

1. INTRODUCTION

1.1. Background and Research Motivation

Probably, everyone must deal with medical wearable devices, such as a pulse oximeter, or some other personal health-related gadgets when he/she starts working at a sleep laboratory. The challenge may require to find a new way to reduce the time of diagnosis for heart-related diseases, sleep disorders, and respiratory-related diseases. There are two main ways to implement this task. First, if it was possible and feasible, software engineering has to be done. Second, If the first way was not feasible, but we had lots of data, either our own or open-sourced, then we could choose the machine learning path and build a data-based model. Figure 1.1 shows the difference between these two ways. Supervised machine learning could be pursued only after data collection and labeling. Also, we would most likely need to validate this work against the standard technology of diagnosis in the industry. Clinical validation and data collection/labeling could be costly. An alternative way of achieving our goal could be done to address both clinical validation and collection/labeling with fewer financial expenses. First, we will need to find an open-source data which will be very similar, in terms of signals/equipment and clinical conditions, to the dataset that we want to label and validate. Second, we need to build a machine/deep learning model using an open-source dataset. Next, we use the trained machine learning model with input data from gold-standard equipment that we are validating against, to label our new clinical dataset. In the last step, we build a new machine learning model using specific signals and labels that were obtained in the previous step. This way, by the end of the project, we will have a validated machine learning model to detect any sleep-related disease from any patient in seconds. Also, we could use this model to help sleep experts in detecting sleep stages from their clinical datasets.

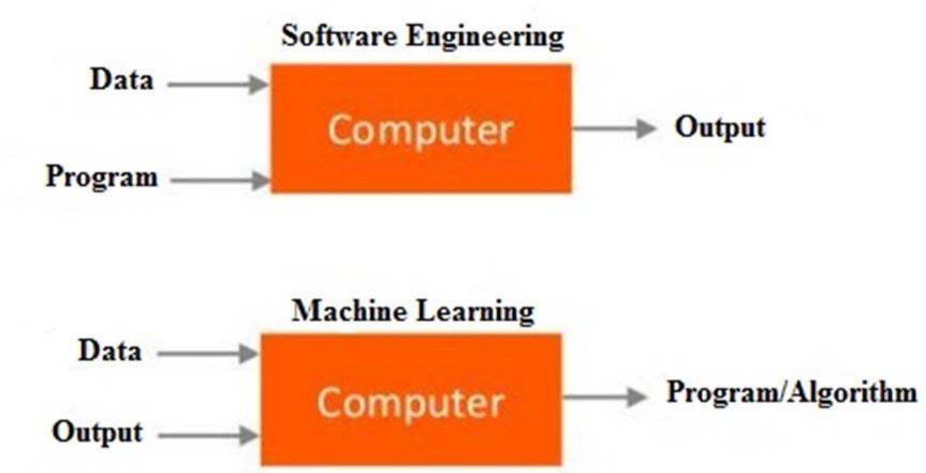

Figure 1.1. Machine learning versus software engineering.

1.2. Problem Statement

Physiological signals are usually difficult in the analysis. In other words, each signal affects human activities separately. Therefore, to study a specific physiological signal, the corresponding software packages are needed. Moreover, an expert person is required to read and search for any abnormality case in that signal, which means that many experts are required. Hence, the difficulties of diagnosis and analyzing those signals are increasing. Therefore, the purpose of this thesis is to provide another solution for analyzing these signals and detecting the sleep-related disease by using deep learning techniques. The provided solution/concept in this thesis is not only related to respiratory diseases; it could also be applied to other diseases like epilepsy. This can be achieved by simply hyper-tuning specific parameters in the model. In this study, we built a model for detecting sleep apnea disease from polysomnography, which is known as a sleep study. Polysomnography is a very laborious process, where the patient must stay a whole night in a laboratory. Sometimes, the patient comes to stay one more night for treatment.

Physiological signals in polysomnography have to be in the EDF/EDF+ format. Format transformation will give errors in analyzing because it makes the signal loses certain specifications in both time and amplitude domains. The solution to avoid these errors is by converting all metadata to the EDF/EDF+ format.

1.3. Contributions of the Thesis

Today, most of the clinical diagnoses are implemented by smart solutions like the internet of things (IoT) and artificial intelligence. Therefore, this thesis shows the ability to use deep learning methods for detecting and validating sleep apnea. The thesis defines and analyzes requirements for applying this approach. As a result, a deep learning model is applied, followed by a validation model for using it on any clinical dataset in this context.

The limitation in the hardware of the sleep laboratory has made most of the physiological signals can't be obtained easily. So that, we accidentally used signals that could be regenerated easily by common sensors. Hence, our work can be used in the unattended sleep test.

The thesis also shows and compares the performance of several deep learning methods. The databases we used also allow users to perform data analysis by using online and offline access methods by visualizing and preprocessing the database on a graphical view.

1.4. Outlines of the Thesis

Chapter 1 presents an overview of the motivation for the thesis, the problems the thesis has to deal with, and the contributions of this thesis. Chapter 2 provides a general discussion of sleep apnea, in which types of the disease and physiological signals of polysomnogram are presented. Chapter 3 gives a literature review of related studies with similarities, gaps, limitations, and concepts. The databases that are used in this thesis are discussed in Chapter 4. There are two types of databases that are presented in the chapter, and both are from Physionet. Algorithms of deep learning that are used are also presented in this chapter, and they are three different algorithms with the hybrid model. In Chapter 5, we discuss the obtained results alongside previous researches. Finally, a synthesis is made, and future works are addressed in Chapter 7.

2. SLEEP APNEA

Sleep apnea is a disorder where breathing pauses and repeatedly starts over through sleeping. Stop breathing occurs when either the airway collapses or the brain cannot successfully send the signal to the breathing muscles. The first case is named Obstructive Sleep Apnea (OSA), and the second case is called Central Sleep Apnea (CSA). In both cases, no air comes in or out of the lung. It often lasts from a few seconds to minutes and can happen about 30 times or more per hour (nhlbi.nih.gov, 2020). There is another case called Hypopnea, which is an irregular respiratory event with at least a 30% decrease in airflow enduring at least 10 seconds and with 4% oxygen desaturation (nhlbi.nih.gov, 2020).

Sleep apnea is difficult to diagnose because it occurs when the patient sleeps. Moreover, sleep apnea cannot be detected by routine tests such as blood tests. It tends to be seen by a relative when the patient intensively snores. However, CSA does not often come with snoring. Therefore, it is challenging to detect sleep apnea in general. Sleep apnea can lead to many severe diseases if it is untreated, such as high blood pressure, heart attack, stroke, heart failure, etc. (mayoclinic.org, 2020). Section 2.1. presents sorts of sleep apnea. Section 2.2. describes the physiological signals that can be used to diagnose sleep apnea.

2.1. Sorts of Sleep Apnea

Sleep apnea is more dangerous than people have thought because it can lead to many other well-known diseases. According to (Hrubos-Strøm et al., 2011), there are 25% of moderately aged Norwegians at high-danger of having OSA. That means sleep apnea is prevalent, and it has to be studied seriously. There are dozens of research papers about it. When searching with “sleep apnea” on Google Scholar, it returns about 643,000 related articles. Sleep apnea is divided into three kinds (mayoclinic.org, 2020):

- **OSA:** It happens when the delicate tissue tumbles to the rear of the throat, which causes the breathing to be obstructed during sleep. If an effort to breathe exists during apnea, then the event is called as obstructive.
- **CSA:** It happens when the cerebrum neglects to impart the signals to the breathing muscles, and therefore it fails to control the breathing cycle. If an effort to breathe doesn't exist during apnea, then the event is called as central.
- **Mixed Sleep Apnea:** It is a mixture of both OSA and CSA. If apnea begins as central apnea, yet close to the end, an effort to breathe has appeared without airflow. Thus, the event is called as mixed.

2.1.1. Central Sleep Apnea

Central sleep apnea syndrome is considered by a cessation or decrease of ventilator effort throughout sleep and is typically connected with oxygen desaturation. CSA happens when the cerebrum can't impart the suitable signals to muscles which control breathing. CSA is not as common as obstructive sleep apnea. CSA is sometimes a consequence of other diseases like heart failure, stroke, or sleep at high altitude (Aurora et al., 2012).

Symptoms: Common symptoms of CSA are awakening suddenly with breathlessness, insomnia, headache in the morning, difficulties to focus on work, hypersomnia, or snoring. The bed-partner can notice that the patient may stop breathing several times during sleep, about 30 times or more per hour, and the patient may snore. However, snoring is not frequent in CSA, and it doesn't generally imply that a patient has apnea (Sleepapnea.org, 2020).

Causes: Many causes make the brain fail to connect with the breathing muscles. One of them is Cheney-Stokes, which is often related to congestive heart failure and is considered by rhythmic cycles. That means gradually increased and decreased breathing (Sleepapnea.org, 2020). Other reasons which cause central sleep apnea are certain medical conditions. For example, if a patient has a cerebrovascular problem or a brain tumor, the breathing system may cause signals to get lost (Wesleypharmacy.com, 2020). Certain drugs such as opium, morphine, or codeine can lead to irregular breathing like increased, decreased, or stopped breathing (Adf.org.au, 2020). Sometimes sleep at high altitudes can cause central sleep apnea, but it is not a big problem because the problem will disappear when moving to a lower altitude (Pagel, Kwiatkowski, and Parnes, 2011).

Complications: Not getting enough sleep can cause cardiovascular diseases. When the oxygen amount in the blood is little, the heart will work hard to deliver blood to tissues. If this lasts for a long time, it may cause heart failure. Likewise, when the heart works hard, the pressure on the vascular increases, which can cause hypertension. Hypertension is one of the most dangerous diseases and can lead to stroke or even death (mayoclinic.org, 2020). Waking up many times at night will make patients tired, and those who do not sleep enough have severe daytime sleepiness, fatigue, and irritability. They cannot focus on work and occasionally fall asleep, not only at work but also when they drive a car.

Diagnostic methods: Usually, a doctor may perform an evaluation based on those symptoms which are mentioned above and decide if there is a need for doing a sleep test. At the treatment center, the patient can be analyzed by overnight observing breathing and other body capacities through sleep. In the polysomnography test, the patient is linked to monitoring devices such as oxygen saturation, arm, and leg movement, breathing patterns, brain activity, lung, and heart. Moreover, an audit by a cardiologist or a specialist who is professional with the nervous system may likewise be important to discover the reasons for CSA (mayoclinic.org, 2020).

Treatments: The first step of treatment is trying to recover from other diseases that cause the CSA. For instance, proper treatment of heart failure may remove CSA. The other way to treat this disease is to use drugs. Some drugs have been used to stimulate breathing for those who have central sleep apnea. For example, acetazolamide can be used to prevent central sleep apnea when sleeping at high altitudes (Pagel et al. 2011). However, other drugs can cause apnea, for example, opioid. The dose of opioid medications needs to be reduced if it causes central sleep apnea. However, drugs are not a significant solution to treat central sleep apnea (Adf.org.au, 2020).

Four physical treatment methods are often used in treating central sleep apnea. The first one is Continuous Positive Airway Pressure (CPAP), which uses gentle air pressure to hold the airways open. This treatment is mainly used for obstructive sleep apnea, where the patient wears a mask on the nose during sleep, but it is also used for central sleep apnea (Pinto and Sharma, 2020).

Another useful method is Bilevel Positive Airway Pressure (BiPAP), which supplies the upper airways with stable and continuous pressure while breathing in and out (Guan et al., 2018). BiPAP provides high pressure when inhaling and low pressure when exhaling. This treatment aims at boosting the weak respiratory pattern of central sleep apnea to normal. Some BiPAP gadgets can detect if there is no air-follow in a few seconds; it will automatically activate the breathing system.

There is a better solution than CPAP and BiPAP, which is called Adaptive Servo Ventilation (ASV). It is designed for the treatment of central sleep apnea and mixed sleep apnea by tracking normal breathing patterns and storing them as a training collection in a database system (Woehrle et al., 2017). When the patient falls asleep, ASV controls the breathing process and compares it with the training set. If there is apnea, ASV may use pressure to control the rhythm of breathing and to avoid respiratory pause.

The last method is using supplemental oxygen, but there are many arguments against this method. The leading case is that the oxygen level may normalize, but the carbon dioxide cannot release. As a result, a signal will be sent to the brain, which causes awakenings, and the sleep will be fragmented (Turnbull et al., 2017).

2.1.2. Obstructive Sleep Apnea

Obstructive sleep apnea syndrome is described by tedious epochs of upper airway obstruction occurring through sleep, typically associated with a decrease in saturation of blood oxygen. It occurs once the throat muscles lounge, and thus, the airways will be closed during sleep. The most noticeable symptom of obstructive sleep apnea is snoring. OSA usually affects older people, but anyone can have this illness. Obstructive sleep apnea is relatively prevalent for overweight people (Kapur et al., 2017). We discussed more in detail about its symptoms, causes, complications, diagnostic methods, and treatments below.

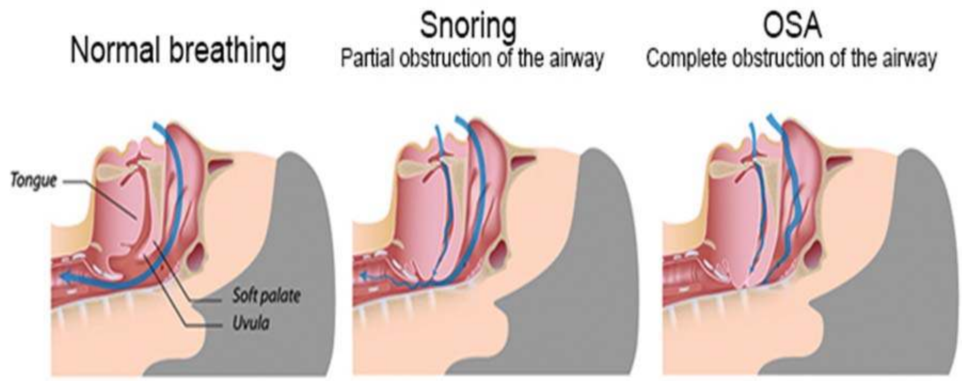

Figure 2.1. Obstructive sleep apnea (360healthcity.com, 2019).

Symptoms: As mentioned above, the most common symptom is loud snoring and sometimes pausing in snoring. Obstructive sleep apnea happens when the throat muscles laze and close the airway. Therefore, when the patient sleeps on his back, the snoring will be loudest, and when he turns on his side, the snoring sound will decrease or even stop. That explains why the patient does not snore every night. It is difficult for doctors to recognize that patients are snoring. However, a family member can simply notice the problem. Another sign is that patients have excessive sleepiness throughout the day, can nod off while working, watching television, or even driving. When the patient does not have normal breathing at night, his brain does not have enough oxygen to function. It causes not only morning headaches, but also memory problems. The other good sign to detect obstructive sleep apnea is waking up with the dried mouth or ulcer throat (Kapur et al., 2017).

Causes: The airway consists of mouth, nose, throat, and windpipe. It always remains open under sleep when there is no sleep apnea. However, the airway becomes narrow when the soft palate and uvula relax. As a result, it becomes difficult to breathe, and the patient might snore. That is called partial obstruction (hypopnea) of the airway. The airway will be blocked if the muscles behind the throat relax too much. As a result, regular breathing is impossible. This situation is called full obstruction of the airway. Blocked in-breath may decrease the oxygen saturation in the blood. Hence, the brain gives a signal to wake up the patient. Then the airway will be opened again. This process usually happens in a short time, and it is often not noticed because shortness of breath may be fixed with one or two deep breaths. Significant causes of obstructive sleep apnea are brain wound, diminished muscle tone, expanded soft tissue around the airway, old age, fat, or structural features that offer ascent to a limited airway (Wikipedia.com, 2020).

The most significant factor is overweight. Nearly half of the people with OSA are overweight because the fat of the upper respiratory can block breathing. The volume of the neck may specify whether someone at risk of OSA or not because the fat neck can close the airway (mayoclinic.org, 2020). Other risk factors are drinking alcohol and smoking. Trenchea et al. (Trenchea et al., 2016) researched this topic and showed that smokers who smoke one pack or more of cigarettes a day are at higher risk of obstructive sleep apnea than those who do not smoke. Other minor but not least causes are sex, age, genetic, or race. Overall, men are twice prone to have OSA than women, and older people have this illness more often than younger.

Complications: Obstructive apnea leads to several complications like cardiovascular diseases. As mentioned in central sleep apnea complications, when the tissues do not get enough oxygen, the heart will beat fast and hard. As a result, the pressure on the heart and vascular will increase. High blood pressure raises the risk of cardiac failure and stroke. Almost half of the obstructive sleep apnea patients have high blood pressure. Besides, sudden death can happen because of the decreased oxygen level in the blood (Health.clevelandclinic.org, 2017).

Moreover, those with obstructive sleep apnea are often capable of having abnormal heart rhythms like atrial fibrillation, medications, and surgery (Aerjournal.com, 2019). In general, the doctor will use some drugs like narcotic analgesics, which can relax the patient's upper airway (mayoclinic.org, 2020). When the patient lies on his back, it will cause severe problems. Therefore, patients with obstructive sleep apnea or related symptoms must inform the physician. Similar to central sleep apnea, daytime drowsiness, fatigue, and irritability are results of repeatedly waking during the night.

Snoring is not only fatigue to the patient but also to those around him. Snoring may be a massive issue for many people. Sleeping with a snoring bed partner sometimes is a nightmare because the people who do not snore can be woken up by the snoring sound. As a result, the partner may get some health problems related to insomnia. In some cases, the partner may leave the sharing room to find a quiet place to sleep, or even it can be a cause for divorce.

Diagnostic methods: The evaluation may usually be performed by observing signs and symptoms. Then some tests need to be done, such as measuring the neck or checking the blood pressure. The medical doctor may examine the back of the throat, mouth, and nose to find out if there are any abnormalities. Further evaluation requires to stay overnight at a sleep center. At the sleep center, breathing and other body functions are checked during the patient's sleep (Kapur et al., 2017).

Sleep tests: A patient is asked to stay at a clinic for monitoring the heart, lung, brain activity, body movements, breathing patterns, and blood oxygen saturation through sleep by using professional devices. Based on the information from those sensors, the physician can diagnose obstructive sleep apnea and its level of severeness. Sometimes a patient has other diseases that also have the same symptoms as obstructive sleep apnea. For example, narcolepsy also causes excessive daytime sleepiness but does not have the same treatment (mayoclinic.org, 2020).

- ***Measuring oxygen:*** The patient does not need to come to a sleep center. Instead, a small sensor that can measure the oxygen level in blood can be used. The patient can stay at home and wear this sensor on his finger to carry out the measurement. It is effortless, consistent, and painless. While the patient sleeps, the sensor collects information about the oxygen level and may store it on a computer or smartphone. Then the data will be sent to the physician. Based on the pattern of oxygen level in the blood, the physician can determine if the patient suffers from obstructive sleep apnea, or the patient may be recommended a sleep test at a sleep center.
- ***Otolaryngology tests:*** This is often performed when a patient snores too loud. As mentioned earlier, when someone snores, it isn't a certain sign that he has obstructive sleep apnea. Anyway, over half of the people who snore have this illness. Hence, examinations on the nose, mouth, throat, palate, and neck need to be carried out.
- ***Cardiopulmonary tests:*** These tests usually relate to the measurements of the airflow, the breath samples, and the heart rate samples. By observing the heart rate and respiration from the abdomen and chest, the physician can identify whether the patient has obstructive sleep apnea or not.

Treatments: Three types of treatments can be applied, which are lifestyle changes, therapies, and surgery, or other procedures (mayoclinic.org, 2020).

- ***Lifestyle changes:*** This is the first treatment to try when OSA is diagnosed. Half of the people who have an obstructive problem are obese, hence trying to lose weight is the first thing to do for the over-weighted patients. Besides weight loss, avoiding drinking alcohol is helpful, since it can also cause this illness. Therefore, the patient should not drink too much alcohol, especially several hours before sleep. An obstructive problem happens when the airway at the throat collapses. Consequently, it is better to lie on the side instead of lying on the back.
- ***Therapies:*** The methods applied for central sleep apnea are also employed for OSA, usually are CPAP, BiPAP, and ASV.
- ***Surgery:*** Surgery aims to get rid of the excess tissue from the throat, which can collapse and block the airway. There are several surgical options: surgical opening in the neck, surgical removal of tissue, jaw surgery, upper airway stimulation, and implants.

2.1.3. Mixed Sleep Apnea

Mixed sleep apnea is also recognized as complex sleep apnea. It is a combination of central and obstructive apnea. It is characterized by researchers at Mayo Clinic (Sciencedaily.com, 2020). Mixed sleep apnea is caused when respiratory control has interfered. Namely, the brain sends a signal to control the breathing system, but the responses of the breathing are not as the brain desired (Sciencedaily.com, 2020). For treatment, CPAP can be used because it can cause chaos on the ventilator control. As a result, obstructive sleep apnea can be healed, but the central sleep apnea still exists because the signals which are sent to the breathing system do not work. New CPAP devices can supply additional carbon dioxide to stabilize the breathing pattern to avoid mixed sleep apnea (Khan and Franco, 2014).

2.2. Physiological Signals of Polysomnogram

Signals, which can be used to find sleep apnea, are divided into two categories. The first group comprises signals that can be simply measured by sensors that are integrated into smart wearable devices. The second group includes complex signals which can only be obtained by professional tools and can be transferred to user devices. The simple signals are the heartbeat and snoring. The complex signals include ECG, EEG, EMG, EOG, Oxygen Saturation (SpO2), Nasal Airflow, and Blood Pressure. Some non-professional devices can also obtain these signals. However, their results are not accurate enough to use for medical diagnosis.

Heartbeat: The standard heart rate of a person varies from 60 to 100 beats per minute, depending on his situation (Carissa Stephens, 2017). Hence, each person has a stable heart rate in that range. However, sleep apnea causes the heart rate to increase and to break the usual pattern. Heart rate is usually easily measured by using a smartwatch.

Snoring: The microphone of a smartphone can measure snoring. The received signal is compared with available data samples. The result of the comparison can be used to determine whether the snoring is related to OSA or not. Snoring is often stopped after a while and starts up again.

EEG: Using only EEG for detecting OSA is not a good idea, because many health problems have similar symptomatology to OSA like epilepsy (Karakis et al., 2012). EEG can be recorded by placing the electrodes along the scalp, which is uncomfortable compared to wearing smart devices. However, EEG is very useful in combination with other signals because it monitors brain activity. Doctors can know whether patients were asleep or awake by monitoring EEG (Karakis et al., 2012). In other words, EEG is considered as an essential signal for detecting OSA.

EMG: People with OSA usually have a longer time for the tongue to be recovered from relaxed. By examining the genioglossus muscle activity, useful data for OSA diagnosis can be collected. For OSA, doctors focus only on the signals from the tongue and genioglossus muscle because they are the leading cause of blocking the airway. Other EMG signals are not considered because they are not relevant (Blumen et al., 2004).

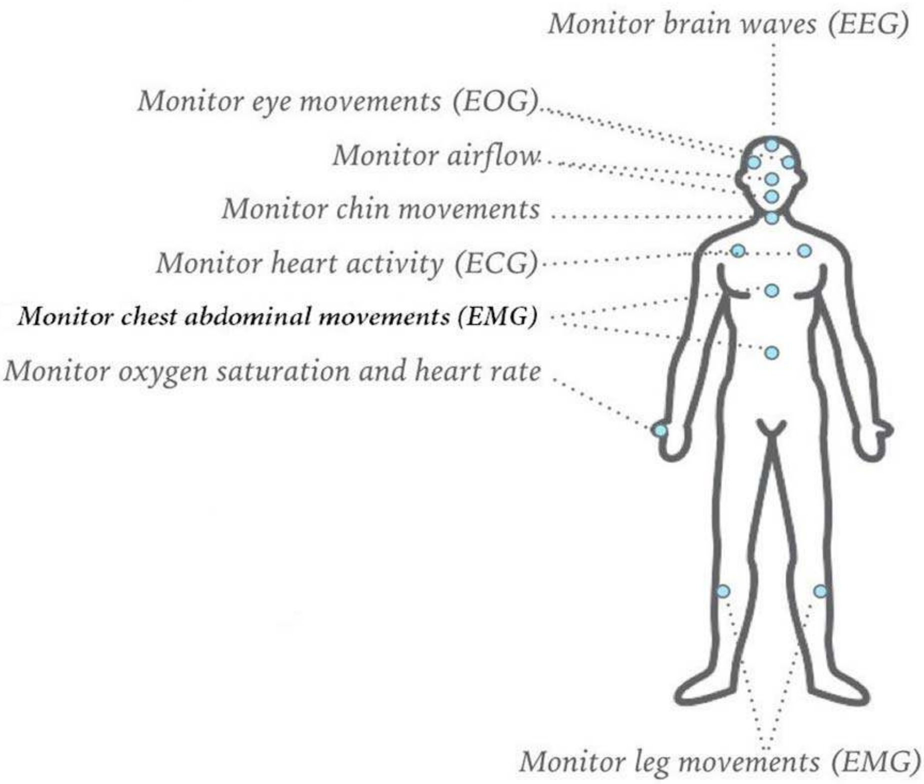

Figure 2.2. Positions of electrodes for polysomnography.

EOG: EOG presents eye movements during sleep. It has two electrodes. The front of the eyes, also recognized as the cornea, is electrically positive while the back of the eyes, also recognized as the retina, is electrically negative. Therefore, it is possible to measure the variation in voltage to determine the movements of the eyes. Rapid eye movement (REM) is a factor in the occurrence of disordered breathing events (Mokhlesi and Punjabi, 2012).

ECG: ECG indicates the heart's electrical activity. For detecting OSA, ECG is the most important signal. By placing the electrodes on the skin, the tiny electrical changes on the skin during each heartbeat can be detected. Once R peak is identified, RR intervals can be calculated, which are the most critical features in the detection of OSA (Almazaydeh, Elleithy, and Faezipour, 2012).

SpO2: Oxygen saturation can be measured by using pulse oximetry. Pulse oximetry is a noninvasive and inexpensive procedure that is used to measure the oxygen levels inside the blood. A clip-like sensor can be placed on the earlobe, nose, or fingers. The level of oxygen in the blood is usually above 95% for healthy people (Wikihow.com, 2020).

Nasal airflow: Airflow sensors can measure the air pressure in the nose when inhaling and exhaling. The sensor can be easily placed, but it is uncomfortable to wear it during sleep. Many studies have reported that it is not a good signal for detecting OSA, but it is still used (Michels et al., 2014).

Blood pressure: Blood pressure can be measured by a simple and noninvasive sensor. This sensor can easily clip on the finger. The average blood pressure with systolic is less than 120 mm Hg, according to the American Heart Association, and with diastolic is less than 80 mm Hg (American Heart Association, 2017). As explained in OSA causes, the pressure is high when the breathing pauses.

3. LITERATURE REVIEWS

Automation of sleep apnea detection has become an interesting topic by many researchers in this field. Most of the previous studies tried to cover it from different aspects. Some of them used more than one database, and others used several physiological signals. They also used various techniques in this task.

The first database that we used in this thesis contains 13 physiological signals, but we used only three signals (chest, abdomen, and oxygen saturation), disregarding electroencephalography and electrooculography signals. Therefore, we are limited to normal and apnea samples (hypopnea, central, obstructive, and mixed apnea). These three signals can be simply obtained from a portable *Fingertip Pulse Oximeter* and *Hexoskin Smart Shirt*. Hence, the cost of the polysomnography study will be reduced by utilizing less equipment and sufficient at the same time. It should be noted that *Hexoskin Smart Shirt* can also provide ECG signals.

In this chapter, we tried to cover almost all studies that are related to the detection of sleep apnea using deep learning techniques from 2008 onwards. We found that most of these findings may contain some exciting areas that must be tested and compared with this thesis.

Novak et al. (Novák, Mucha, and Al-Ani, 2008) demonstrated the use of LSTM with a single ECG signal. The apnea diagnosis was carried out on the characteristics derived from the Heart Rate Variability (HRV) tests. LSTM was chosen to report the time-based dependency in HRV data because one of the critical aids of LSTM is the capability to use the prior situation.

Pathinarupothi et al. (Rahul K. Pathinarupothi et al., 2017) and (Rahul Krishnan Pathinarupothi et al., 2017) suggested using LSTM to detect the OSA severity by utilizing one feature, which is Instantaneous Heart Rate (IHR). Then, by adding an extra feature, which is the oxygen saturation (SpO2). This technique is highly preferable by physicians to apply to their patients.

Haidar et al. (Haidar, Koprinska, and Jeffries, 2017) (McCloskey et al., 2018) and (Haidar et al., 2018) reported a CNN architecture to detect apnea with different trials. First, they used the Nasal Airway signal to equate CNN to Support Vector Machine (SVM). Later, the 2D spectrogram images of the Nasal Airflow signal with the raw 1D Nasal Airway signal were used to apply a different strategy. Lastly, they obtained a consistent result by using three separate signals, namely Nasal Airflow, Abdominal, and Thoracic.

Nesaragi et al. (Nesaragi et al., 2018) proposed an LSTM model consisting of two stages. They neglected the SpO2 signal. They discovered the strong point of using instantaneous frequency and spectral entropy features for the detection of arousals. They promised to improve performance.

Biswal et al. (Biswal et al., 2018) provided an excellent generalization by training the RCNN model on MGH data, whereas the testing was on SHHS data. We can notice that accuracy has been decreased using this approach.

Urtnasan et al. (Urtnasan, Park, and Lee, 2018) (Banluesombatkul, Rakthanmanon, and Wilaiprasitporn, 2018) (Urtnasan, Park, and Lee, 2018) and (Erdenebayar et al., 2019) used a single ECG signal with CNN, LSTM, and GRU techniques. They recorded high performance but, they could have proved that by using multi-signal.

Li et al. (Li et al., 2018) also proposed a novel approach for improving accuracy by integrating the Hidden Markov Model (HMM) with Deep Neural Network (DNN). In addition, the fusion decision algorithm has been adopted to boost overall performance.

Wang et al. (Wang et al., 2019) developed an improved LeNet-5 convolutional neural network with ECG segments. According to existing methods of machine learning, the use of transfer learning models is proved to be useful and promising.

Islam et al. (Islam et al., 2019) proposed a novel way of detecting sleep apnea using 3D scans. They used pre-defined models in transfer learning to beat the limitation of the small data set. Their preceding work has shown the connection between facial morphology and OSA.

Finally, Mahmud et al. (Mahmud et al., 2020) pointed out that the lack of data can cause low performance. They used EEG signals composed of only seven different apnea patients. Table 3.1 shows those results in terms of databases, models, classes, and accuracy.

Table 3.1**.** A review of our results with those reported in other studies**.**

No	Authors	Databases	Methods	Classes	Accuracy
1	Novak et al. (Novák, Mucha, and Al-Ani, 2008)	Apnea-ECG	LSTM	A/N	82.1
2	Kaguara et al. (Kaguara, Myoung Nam, and Reddy, 2014)	Apnea-ECG	Autoencoder	A/N	91
3	Rahul K et al. (Rahul K. Pathinarupothi et al., 2017)	Apnea-ECG	LSTM	A/N	85
4	Rahul Krishnan et al. (Rahul Krishnan Pathinarupothi et al., 2017)	Apnea-ECG MIT-BIH (PhysioNet, 2005)	LSTM	A/N	100 99.9
5	Cheng et al. (Cheng et al., 2017)	Apnea-ECG	LSTM	OA/N	97.8
6	Mostafa et al. (Mostafa et al., 2017)	Apnea-ECG UCD (Physionet 2008)	DBN	A/N A/N	97.64 85.26
7	Haidar et al. (Haidar, Koprinska, and Jeffries, 2017)	MESA (Dean et al., 2016) (G.-Q. Zhang et al., 2018)	CNN	A/N	75
8	McCloskey et al. (McCloskey et al., 2018)	MESA	CNN	OA/H/N	79.8

Table 3.1. continuous

9	Li et al. (Li et al., 2018)	Apnea-ECG	DF-DNN	OA/N	84.7
10	Lakhan et al. (Lakhan et al. 2018)	MrOS (Blank et al., 2005)	DNN	Mi/Mo/Se/N	63.70
11	Nesaragi et al. (Nesaragi et al., 2018)	YSYW	LSTM	Arousal/Non-Arousal	62.4 (AUROC)
12	Choi et al. (Choi et al., 2018)	MESA SNUH	CNN	A/H/N	96.6
13	Haidar et al. (Haidar et al., 2018)	MESA	CNN	OA/H/N	83.5
14	Urtnasan et al. (Urtnasan, Park, Joo, et al. 2018)	Samsung Medical Center	CNN	OA/N	96
15	Zhang et al. (L. Zhang et al., 2018)	SHHS (Quan et al., 1997)	CNN	A/H/N	88.5
16	Cen et al. (Cen et al., 2018)	UCD	CNN	A/H/N	79.61
17	Alsalamah et al. (Alsalamah, Amin, and Palade, 2018)	Apnea-ECG	DNN	A/H/N	92.7
18	De Falco et al. (De Falco et al., 2018)	Apnea-ECG	DNN	OA/N	68.36
19	Pomprapa et al. (Pomprapa et al., 2018)	Contactless Multisensor	RCNN	A/N	90
20	Biswal et al. (Biswal et al., 2018)	MGH (PhysioNet, 2002) SHHS	RCNN	A/N	88.2 80.2
21	Banluesombatkul et al. (Banluesombatkul, Rakthanmanon, and Wilaiprasitporn, 2018)	MrOS	RCNN	OA/N	79.45
22	Dey et al. (Dey, Chaudhuri, and Munshi, 2018)	Apnea-ECG	CNN	OA/N	98.61

Table 3.1. continuous

23	Islam et al. (Islam et al., 2019)	3D Scans	VGG Face VGG19 AlexNet	OA/N	67.42 57.14 59.37
24	Van Steenkiste et al. (Van Steenkiste et al., 2019)	SHHS	LSTM F-LSTM	OA/CA/H/N Mi/Mo/Se/N	77.2 74.7
25	Wang et al. (Wang et al., 2019)	Apnea-ECG UCD	LeNet-5 (modified)	OA/N OA/N	87.6 71.8
26	Chaw et al. (Chaw, Kamolphiwong, and Wongsritrang, 2019)	Songklanagarind Hospital Thailand	CNN	A/H/N	91.3
27	Erdenebayar et al. (Erdenebayar et al., 2019)	Samsung Medical Center	DNN 1D CNN 2D CNN RNN LSTM GRU	A/H/N	93.1 98.5 95.5 85.4 98 99
28	Mahmud et al (Mahmud et al., 2020)	UCD	DNN	A/N	87.67

N: Normal, A: Apnea, H: Hypopnea, OA: Obstructive Apnea, CA: Central Apnea, Mi: Mild, Mo: Moderate, Se: Severe

4. MATERIAL AND METHODS

In this chapter, we covered the used databases, the deep learning algorithms, and the platform of the training and test. Lastly, we presented the metrics, which were used to measure the performance of algorithms.

4.1. Sleep Apnea Databases

4.1.1. You Snooze You Win Database (Challenge 2018)

The Massachusetts General Hospital (MGH) Computational Clinical Neurophysiology Laboratory (CCNL) and the Clinical Data Animation Center (CDAC) provided this database (Ghassemi et al., 2018). It includes 1,985 patients for the detection of sleep disorders obtained in an MGH sleep laboratory. They split the database into 994 folders as a training set and 989 folders as a test set.

The patients' sleep stages were labeled by clinical staff at the MGH, according to the American Academy of Sleep Medicine (AASM) guideline. More precisely, six sleep stages, which are wakefulness, N1, N2, N3, Rapid Eye Movement (REM), and Undefined, were labeled in 30-second intervals.

The database was also labeled by certified sleep technologists at the MGH according to the existence of arousals. These derived arousals were either classified as: Respiratory Effort-Related Arousals (RERA), spontaneous arousals, hypoventilations, bruxisms, hypopneas, apneas (central, obstructive, and mixed), vocalizations, snores, periodic leg movements, breathing Cheyne-Stokes, or partial airway obstructions.

Many physiological signals, consisting of electroencephalography (EEG), electrooculography (EOG), electromyography (EMG), electrocardiography (ECG), and oxygen saturation (SaO2), were taken from the patients when they slept at night.

The database includes two directories (training and testing), each approximately 135 GB in volume. Each directory contains one subfolder per patient (e.g., training/tr06-0475). Every subfolder includes signal, header, and arousal files such as the example below:

- **tr06-0475.mat:** It includes signal data.
- **tr06-0475.hea:** It displays the signal data format.
- **tr06-0475.arousal:** It has labels for the arousal and the sleep stages.
- **tr06-0475-arousal.mat:** It contains three distinct values (+1, 0, -1) of which:
    - **+1:** It specifies areas that are arousal.
    - **0:** It specifies areas that are not arousal.
    - **-1:** It specifies areas that are not labeled.

Test sets are unlabeled. Therefore, we were unable to use them in the testing phase. Records in the database were taken from different persons in both sex and age. The following table shows an example from the training set.

Table 4.1**.** A subset of the training set, which shows the name, sex, and age.

Record	**Sex**	**Age**
tr03-0005	M	59
tr03-0029	M	55
tr03-0052	F	46
tr03-0061	M	43
tr03-0078	F	47
tr03-0079	M	29
tr03-0083	M	63
tr03-0086	M	54
tr03-0087	M	24
tr03-0092	M	43
tr03-0100	M	67
tr03-0103	M	52
tr03-0134	M	49
tr03-0135	F	46
tr03-0141	F	68

Class distribution for apneas and normal samples is unbalanced, where 19% of the total samples are apnea. The problem was fixed by preprocessing the data. The process is given in section 5.1.

4.1.2. Apnea-ECG Database (Challenge 2000)

This challenge's database (Penzel et al., 2000) includes 70 records, divided into two equal parts, 35 records as a training set (a01 to a20, b01 to b05, and c01 to c10), and 35 records as a test set (x01 to x35). Only the learning sets were annotated. Each record includes signals, headers, and other data, including:

- **a01.dat:** It provides the ECG signal data.
- **a01.hea:** It describes the format of the signal data.
- **a01.apn:** It contains labels for every 60 seconds of each record, referring to the appearance or nonappearance of apnea at that time.
- **a01.qrs:** It is given for the suitability of those who do not want to utilize their QRS detectors.
- **a01.xws**: It views any record without downloading it.

Besides, eight records (a01, b01, c01, a02, c02, a03, c03, and a04) are followed by four additional signals (CHEST, ABD, Nasal Airflow, and SpO2). Three of them were used for further testing of our models.

4.2. Deep Learning Techniques

In this section, four techniques of deep learning were presented with equations and significant figures. The first two techniques are related to recurrent neural networks (RNNs). The third technique is related to convolutional neural networks (CNNs). The fourth technique is a hybrid model consisting of RNNs and CNNs.

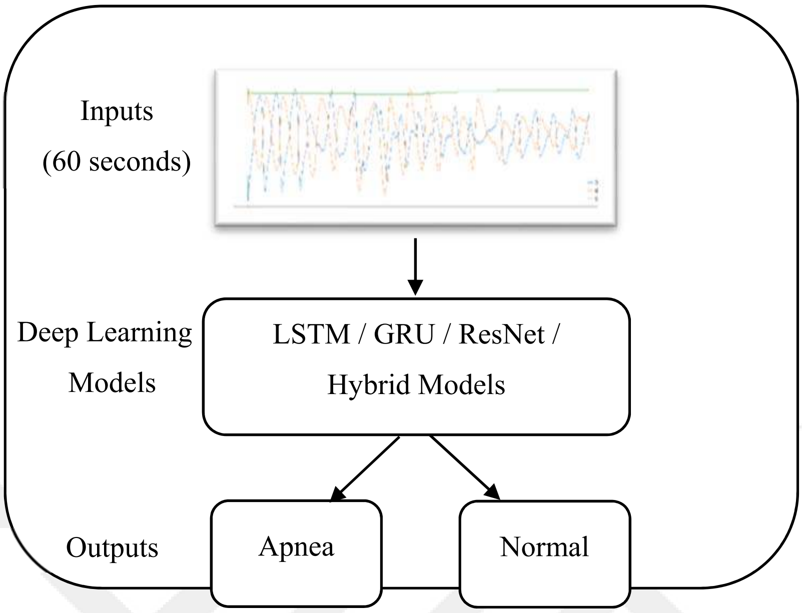

Figure 4.1. Sleep apnea events detection approach

4.2.1. Long Short-Term Memory (LSTM)

Hochreiter and Schmidhuber (Hochreiter and Schmidhuber, 1997) initially suggested the expression of Long Short Term Memory (LSTM) in 1997. LSTM is a kind of RNNs, and it became very popular in recent years due to its significant performance and in solving the vanishing gradient problem. LSTM overcomes that by imposing fixed error flow. LSTM clearly learns when to save the information and when to retrieve it using gradient descent.

LSTM is the most complex recurrent unit. The difficulty of mathematics does not cause that complexity. It is in terms of the high number of components.

LSTM has distinct elements in the recurrent hidden layer, called memory blocks, whereas RNN does not have them. Such blocks provide memory cells with self-connections to preserve the network's time-based state as well as different multiplicative units called gates to alter information flow. Each memory block within the interior architecture consists of three types of gates, as shown in Figure 4.2, which are specifically:

- **Input gate:** It dominates the quantities that go into the cell of the new value.
- **Output gate:** It takes into account the input at time t, the prior hidden state, and the present value of the cell.
- **Forget gate:** It dominates the ex-cell value quantities that go into the present cell value.

The memory cell in an LSTM network works as a single unit within the hidden layer of traditional networks. The formulas are given in the following equations:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{4.1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4.2}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{4.3}$$

$$\widetilde{C_t} = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4.4}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \tag{4.5}$$

$$h_t = o_t * \tanh(C_t) \tag{4.6}$$

Where:

f_t, i_t, o_t: Forget, input, and output gates.

h_t, C_t: Hidden layer vectors.

x_t: Output vector.

b_f, b_i, b_o, b_c: Bias vector.

W_f, W_i, W_o, W_c: Parameter matrices.

σ, *tanh*: Activation functions.

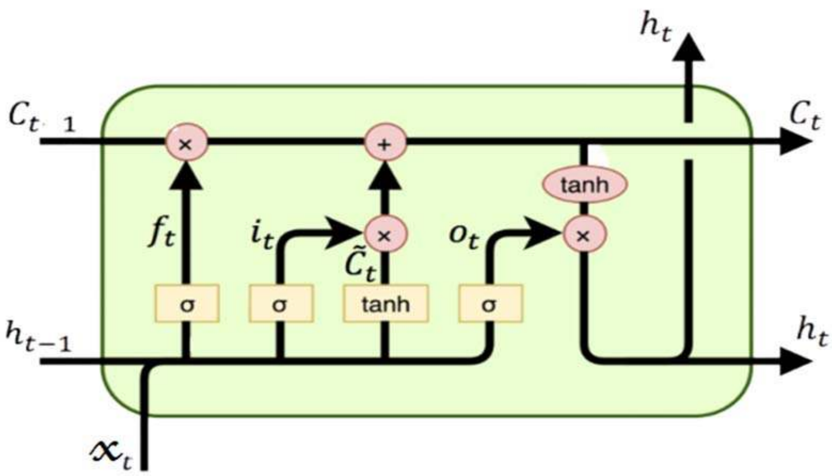

Figure 4.2. LSTM memory block, where f_t, i_t, and o_t represent the forget, input, and output gates, respectively. c_{t-1} and c_t represent the memory cell and the content of the new memory cell (Dprogrammer.org, 2020).

4.2.2. Gated Recurrent Unit (GRU)

Gated recurrent unit (GRU) was announced in 2014 by Cho et al. (Cho et al., 2014). GRU is a simple form of LSTM. It joins many of the same concepts, but it has a much smaller set of parameters so that it can be trained more quickly at a sustained hidden layer size. Recent research has also shown that the accuracy between the LSTM and GRU is comparable and even better with the GRU in some cases. GRU includes the same internal structure of LSTM, but the memory cell consists of two gates rather than three, as in figure 4.3. which are namely:

- **Update gate:** It also compares how much of the hidden value of the previous candidate and how much of the hidden value of the current candidate is combined to get the new hidden value.
- **Reset gate:** It has the same physical shape as the update gate, and all its weights are equal, but its position in the memory block is different. It controls how much of the previous hidden state is considered when the

new candidate hidden value is created. In other words, it can “reset” the hidden value.

$$z_t = \sigma(W_z . [h_{t-1}, x_t] + b_z) \tag{4.7}$$

$$r_t = \sigma(W_r . [h_{t-1}, x_t] + b_r) \tag{4.8}$$

$$\widetilde{h_t} = tanh(W_h . [r_t * h_{t-1}, x_t] + b_h) \tag{4.9}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \widetilde{h_t} \tag{4.10}$$

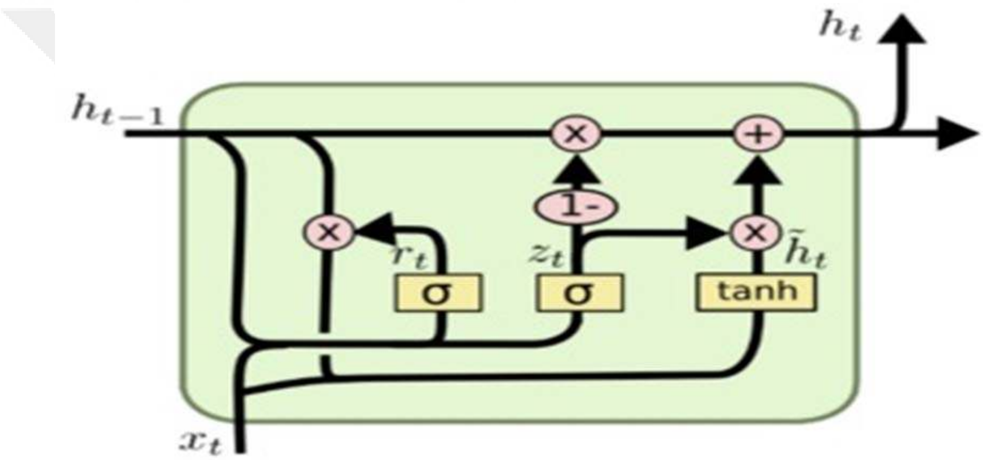

Figure 4.3. GRU memory block, where r_t and z_t denote the reset and update gates, and h_t and $\widetilde{h_t}$ denote the activation and the candidate activation (Dprogrammer.org, 2020).

Where:

h_t: Hidden layer vectors.

z_t, r_t: Update and reset gates.

x_t: Output vectors.

b_z, b_i, b_r, b_h: Bias vectors.

W_z, W_r, W_h: Parameter matrices.

σ, *tanh*: Activation functions.

4.2.3. Residual Network (ResNet)

Residual network (ResNet) is a pre-trained deep neural network model. Pre-trained models are very handy. We can use them in many tasks like prediction and feature extraction, or we can fine-tune them to a specific case. The motivation behind transfer learning is built on the fact that people can wisely use the knowledge acquired earlier from different missions or issues to handle new problems quicker or with improved solutions (Pan and Yang, 2009). Figure 4.4 shows the difference between machine learning and transfer learning.

We used the ResNet50 model as the pre-trained model for sleep apnea events detection. We eliminated the predicting layer from its architecture and substituted it with our predicting layer. Weights of the first few layers are untouched or updated during the training because the first few layers save general information like curves and edges, which are also related to our task. Instead, we made the network to emphasis on learning specific features in the subsequent layers.

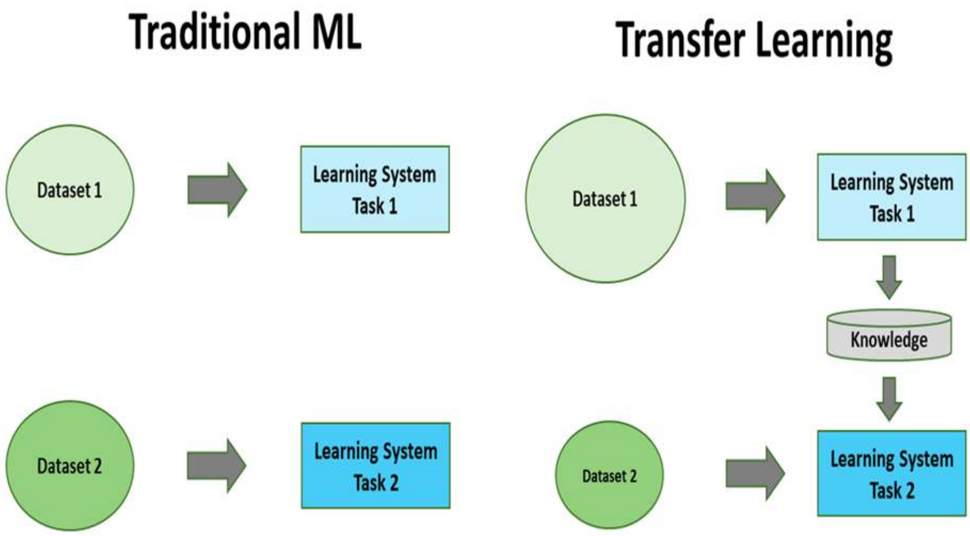

Figure 4. 4. The difference between machine learning and transfer learning.

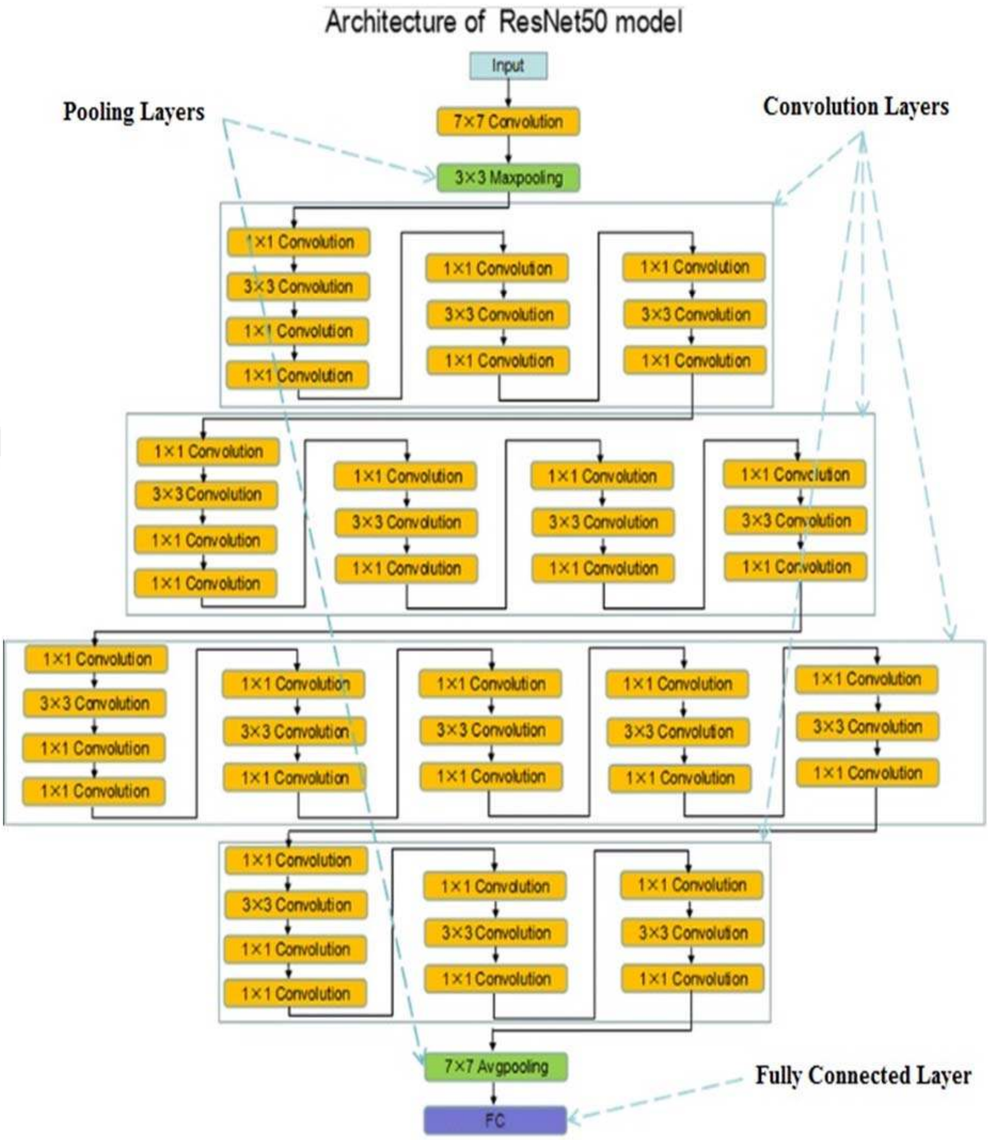

Figure 4. 5. The architecture of the ResNet50 model (Peng et al., 2020).

4.2.4. A Hybrid Model

This model consists of the ResNet50 architecture with two RNN layers. One GRU layer comes after the input layer, and one bidirectional LSTM layer comes before the output layer. The full detailed architecture of this model is shown in Figure 4.6.

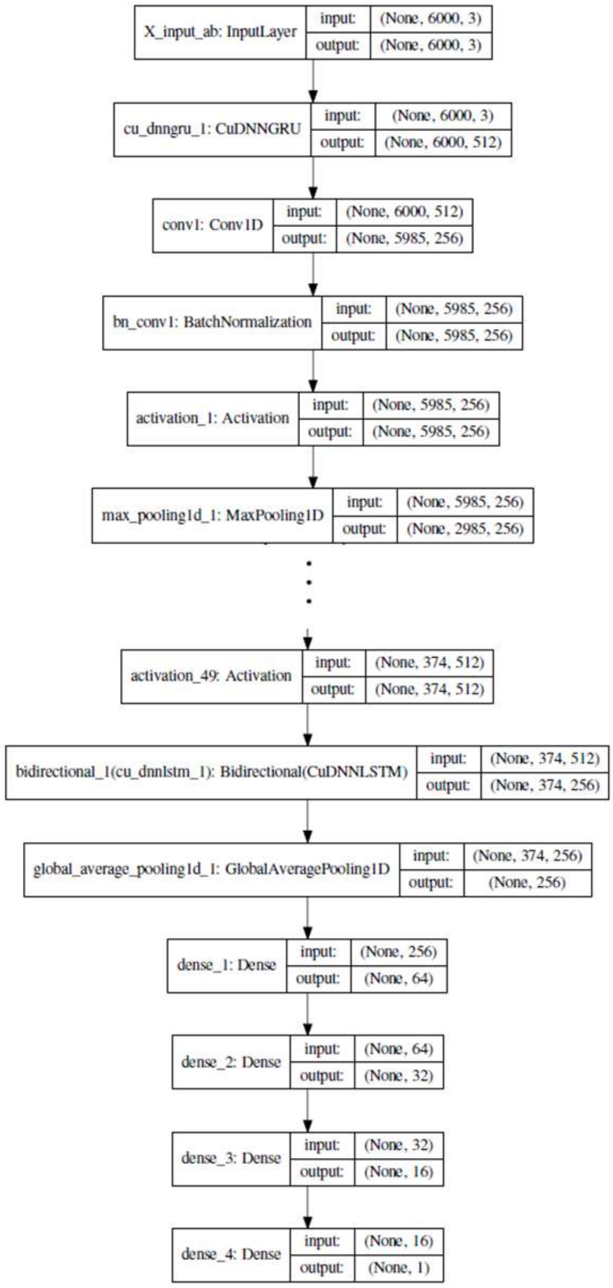

Figure 4. 6. The architecture of the hybrid model.

4.3. Training and Testing Platform

We used Google Colaboratory (Bisong, 2019), which is a development environment in Python with high capabilities. It is a free cloud service based on jupyter notebooks for creating projects in the machine and deep learning. Colab is also combined with Google Drive, so users can simply share and move their notebooks with the related materials onto their accounts.

Colab offers a runtime fully configured for deep learning with free access to a robust GPU. The only limitation of Colab is the interruption of the runtime, which stops after 12 hours of continuous working. Sometimes, it disconnects after one hour of inactivity. That is why it is so important to speed up the rerun time of our runtime. We also noticed that Colab gives 12 GB of RAM and increases it up to 25 GB after runtime if required. The speed of the internet in Colab is up to 120 MB/s. So, we downloaded the training set of the YSYW database directly into Colab. We suffered from RAM crashing because of the big size of the data. Therefore, we created several text files to download the data in multi-stages using the wget command line. Colab has nearly all the necessary Python libraries pre-installed like TensorFlow, Keras, Pandas, Matplotlib, and Numpy.

4.4. Performance Assessments

We need to find out the accuracy of the model after the training phase by the deep learning algorithms. So, the test set is used to evaluate the model. The test set includes samples that have never been seen before by the algorithm. If the model performs well in predicting, it can be assumed that the model is generalizing well.

Generally, experts in machine learning use various formal metrics and methods to assess model efficiency. The most widely used metrics and tools for assessing the classification model are:

Confusion Matrix: It is a table that summarizes how successful the classification model is at predicting examples belonging to various classes. One dimension of the confusion matrix is the label predicted by the model, and the other dimension is the label itself.

Table 4.2. The shape of the binary confusion matrix.

		Predicted	
		Normal	***Apnea***
Actual	***Normal***	TN	FP
	Apnea	FN	TP

Where:

True Positive (TP): Positive instances truly predicted

True Negative (TN): Negative instances truly predicted

False Positive (FP): Positive instances falsely predicted

False Negative (FN): Negative instances falsely predicted

Accuracy: It is the ratio of instances truly predicted, divided by the truly and falsely predicted instances.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{4.11}$$

Precision: It is the ratio of positive instances truly predicted, divided by positive instances truly and falsely predicted.

$$Precision = \frac{TP}{(TP+FP)} \tag{4.12}$$

Recall or Sensitivity: It is the ratio of positive instances truly predicted, divided by positive and negative instances truly and falsely predicted.

$$Recall = \frac{TP}{(TP+FN)} \tag{4.13}$$

F-Score: It is a weighted average of precision and recall. So, this ranking takes into account all false positives and false negatives. F-score is typically more helpful than accuracy, particularly when the distribution of classes is unequal. Accuracy is better if the cost of false positives and false negatives is close. If the importance of false positives and false negatives is different, then it is better to look at accuracy, recall, and f-score together.

$$F - score = \frac{(2 * Recall * Precision)}{(Recall + Precision)} \tag{4.14}$$

5. EXPERIMENTAL RESULTS

The main task of this thesis is to use deep learning techniques to detect apnea events (hypopnea, obstructive, central, or mixed) by handling several physiological signals. The essential benefits of our work are the following outlines:

- Detection of sleep apnea events from the YSYW database test set
- Detection of sleep apnea events from the Apnea-ECG database test set without training.
- Detection of sleep apnea events from any other database has the same corresponding signals.

5.1. Data Preprocessing

Some standard Python packages were used to access PhysioNet (Goldberger et al., 2000). They were also used to display and prepare signals for apnea events detection. Firstly, we worked on the YSYW database; then, we used the Apnea-ECG database for testing the trained model, as we mentioned that in Chapter 4. There are 994 patients in training set with labels for sleep apnea events, sleep stages, and arousals. The test set has 989 patients without labels.

Practically, it is not necessary to download the training set for data preprocessing because we used the WFDB API package (Xie and Dubiel, 2016) in Python for remote access. After importing the necessary libraries, setting up plot parameters for a better view, and setting up the exact paths for the database, we got 13 physiological signals in each PSG. Signals were measured in microvolts, excluding oxygen saturation (SpO2), which was measured as a percentage. We selected three signals (ABD, CHEST, and SpO2) and dropped the rest from our estimations. An apnea events dictionary was created, containing only apnea labels.

On the other hand, all labels for apnea/stages were pulled from PhysioNet and combined with those three signals. They were loaded into a data frame, and then the apnea events dictionary that we created was mapped with the same data frame. Hence, other unused labels (sleep stages) in that dictionary were shown as missing values. Therefore, they were replaced by zero. Labels are finally encoded into the one-hot array. Figure 5.1 shows the raw targeted signals before preprocessing, and Figure 5.2 shows the CHEST signal annotated by an obstructive apnea.

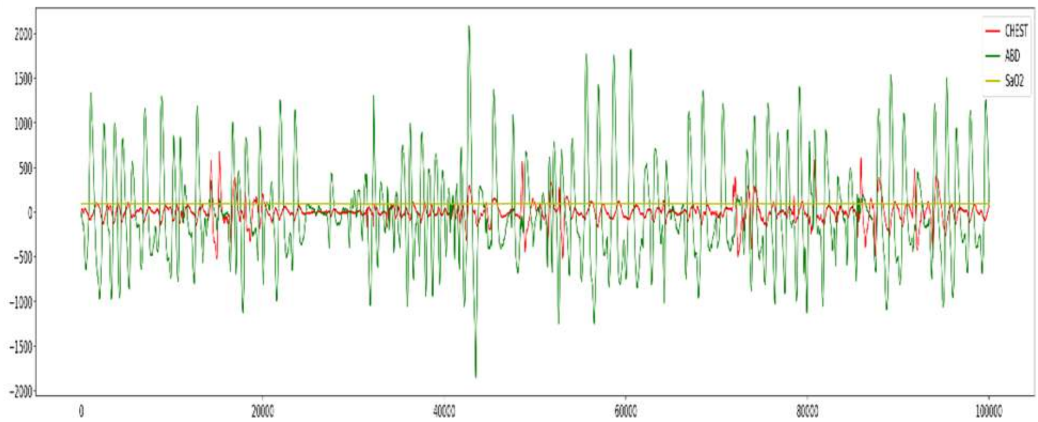

Figure 5. 1. Unscaled signals (before preprocessing).

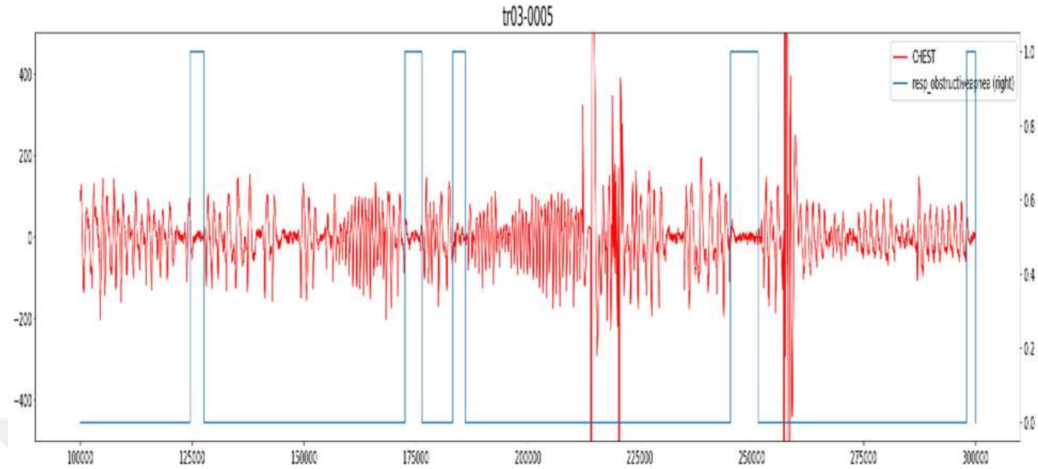

Figure 5. 2. CHEST signal of patient tr03-005 annotated by an obstructive apnea.

Signals where SpO2 is below 50% were removed because there is a possibility that the connection with devices would be lost. The sampling frequency was resampled to 100 Hz because 100 samples per second are enough to use. The imbalanced problem between apnea and normal classes was fixed by under-sampling. Finally, we reshaped signals into intervals with a fixed length. We applied the minimum-maximum method to scale data between 0 and 1 for normalization according to Equation 5.1. Figure 5.3 shows the final shape of signals before using them for the task of apnea events detection. Table 5.1 summarizes all these steps in concise.

$$X_{new} = \frac{X_{old} - X_{min}}{X_{max} - X_{min}} \tag{5.1}$$

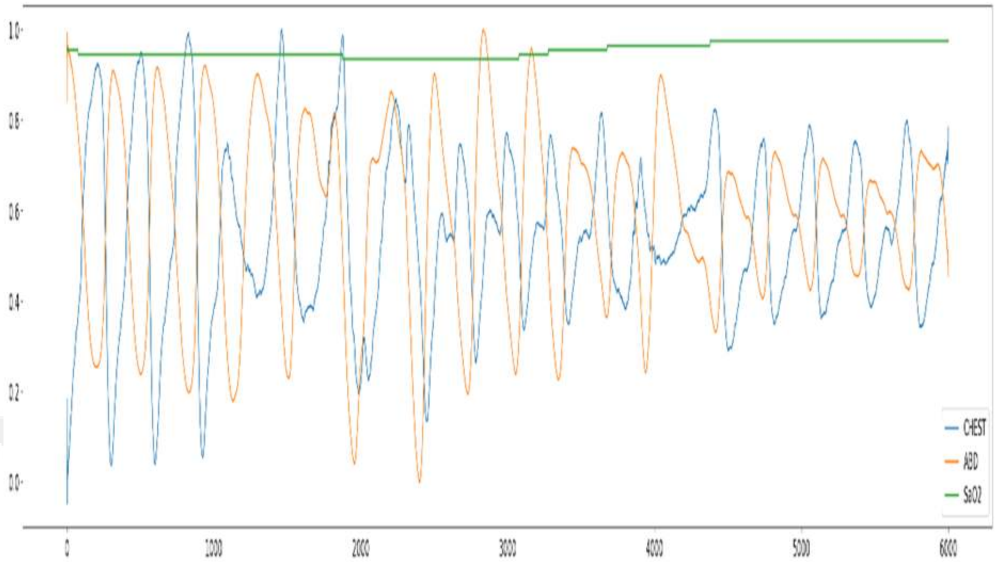

Figure 5. 3. Scaled signals (after preprocessing).

Table 5. 1. Summary of data preprocessing.

No	Procedures
1	Removing tr03-0394, tr07-0016 and tr07-0023 (no labels for apnea)
2	Selecting CHEST, ABD, and SpO2 signals and dropping the rest
3	Pulling labels for only sleep apnea events and combining them with signals
4	Other unused labels (sleep stages) were replaced by zero
5	Encoding labels to the one-hot array
6	Removing signals where SpO2 is below 50%
7	Resampling the sampling frequency to 100 Hz instead of 200 Hz
8	Under-sampling the ratio between classes (Apnea/Normal)
9	Normalization signals' values between 0 and 1

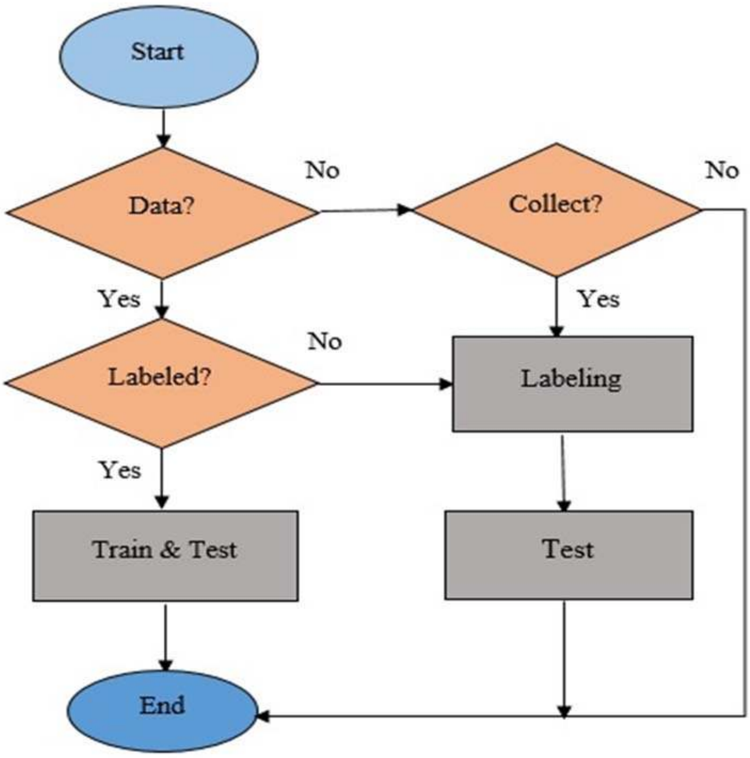

Figure 5. 4. Sleep apnea events detection workflow (supervised learning).

5.2. Deep Learning Techniques

Four techniques were chosen for comparison according to their performances. The following flow chart in Figure 5.4 shows the algorithm of our method for the detection of sleep apnea events.

We extracted 39209 windows from the test set of the YSYW database, which represents the last 91 patients of the training set. Besides, we also extracted 2849 windows from the test set of the Apnea-ECG database, which represent only eight patients who have signals of ABD, CHEST, and SpO2. These windows were gathered from every patient's data. Each window was formatted as a 1 × 6000. The mathematical meaning of this representation is that we multiplied the window's size, which is 60 seconds, by the sampling frequency. The amplitude of signals was normalized between 0 and 1. The preprocessing steps were explained in section 5.1.

5.2.1. Long Short-Term Memory (LSTMs) Results

The LSTM model used is composed of four layers of LSTM. Each layer has either 30, 60, 120, or 200 memory cells. After each LSTM layer, there are batch normalization and dropout to avoid overfitting. For classification, one fully connected layer was used at the output layer. The confusion matrices are shown in Figure 5.5 and Figure 5.6.

Table 5.2 consists of the evaluation metrics of the LSTM model, where accuracy is the overall accuracy (Apnea/Normal), whereas precision, recall, and F1 score belong to the apnea class.

Table 5. 2. Evaluation metrics of the LSTM model.

No	Database	Accuracy	Precision	Recall	F1 Score
1	YSYW	0.87	0.76	0.77	0.76
2	Apnea-ECG	0.87	0.87	0.89	0.88

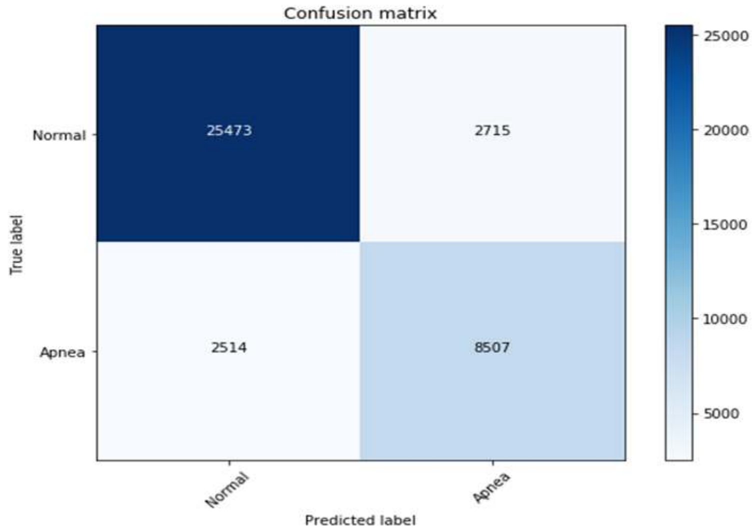

Figure 5.5. Confusion matrix of the LSTM model applied to the YSYW database. Where TP: 8507, TN: 25473, FP: 2715, FN: 2514. Normal: 0 and Apnea: 1

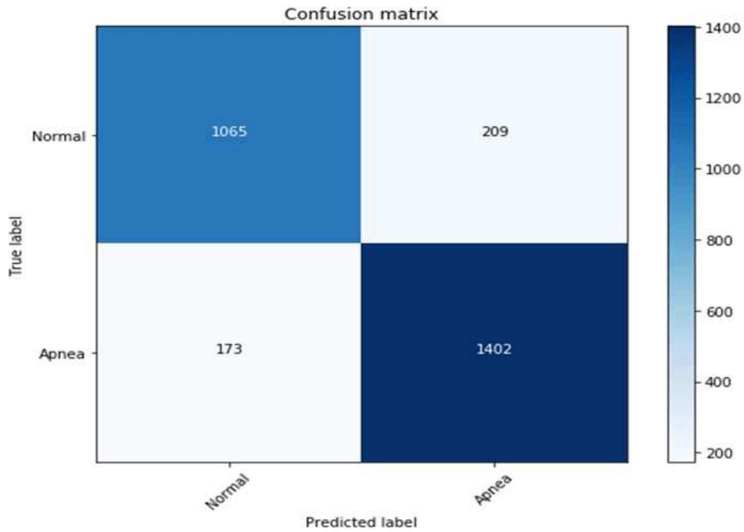

Figure 5 6. Confusion matrix of the LSTM model applied to the Apnea-ECG database. Where TP: 1402, TN: 1065, FP: 209, FN: 173. Normal: 0 and Apnea: 1

5.2.2. Gated Recurrent Units (GRUs) Results

The utilized GRU model consists of four layers of GRU. Each layer has either 30, 60, 120, or 200 memory cells. After each GRU layer, there are batch normalization and dropout to avoid overfitting. For classification, one fully connected layer was used at the output layer. The confusion matrices are shown in Figure 5.7 and Figure 5.8.

Table 5.3 consists of the evaluation metrics of the GRU model, where accuracy is the overall accuracy (Apnea/Normal), whereas precision, recall, and F1 score belong to apnea class detection.

Table 5. 3. Evaluation metrics of the GRU model.

No	Database	Accuracy	Precision	Recall	F1-Score
1	YSYW	0.92	0.85	0.87	0.86
2	Apnea-ECG	0.91	0.89	0.93	0.91

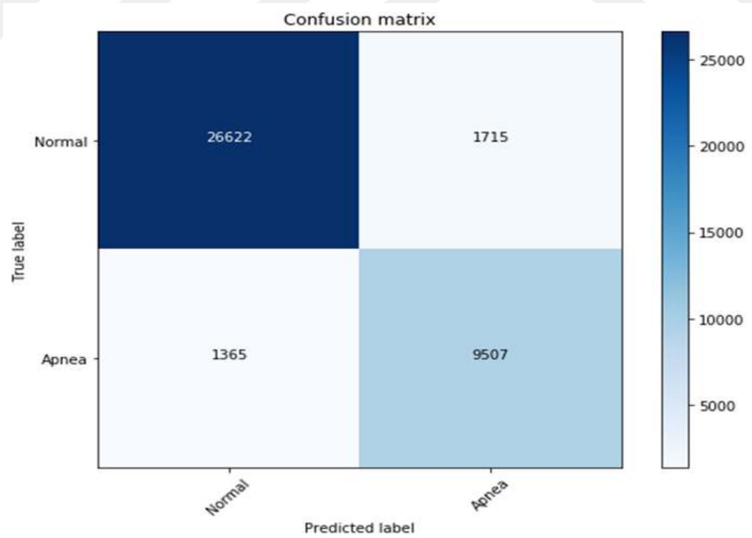

Figure 5.7. Confusion matrix of the GRU model applied to the YSYW database. Where TP: 9507, TN: 26622, FP: 1715, FN: 1365. Normal: 0 and Apnea: 1

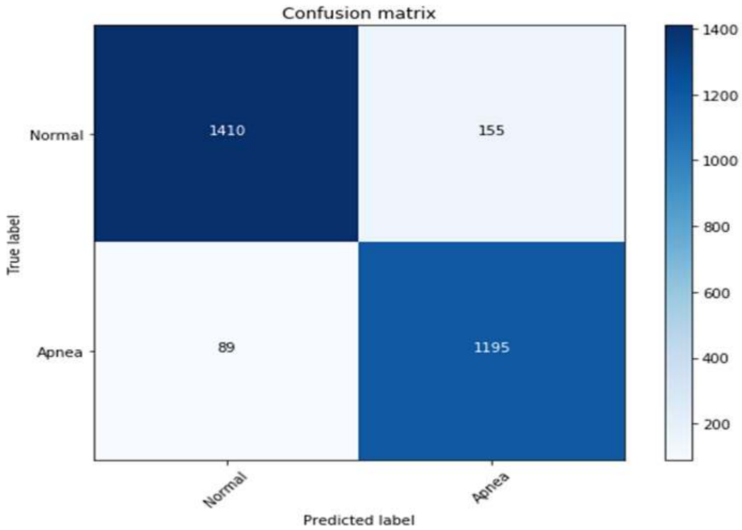

Figure 5.8. Confusion matrix of the GRU model applied to the Apnea-ECG database. Where TP: 1195, TN: 1410, FP: 155, FN: 89. Normal: 0 and Apnea: 1

5.2.3. Residual Networks (ResNets) Results

The utilized ResNet model consists of 48 convolution layers, one MaxPool layer, and one AveragePool layer. Convolution layers with different kernel size and activation functions are connected. After that, there is an average pool layer followed by a fully connected layer containing 1000 nodes with softmax function at the end. The full architecture of ResNet50 is shown in figure 4.4. The confusion matrices of this model are shown in Figure 5.9 and Figure 5.10.

Table 5.4 consists of the evaluation metrics of the ResNet model, where accuracy is the overall accuracy (Apnea/Normal), whereas precision, recall, and F1 score belong to apnea class detection.

Table 5. 4. Evaluation metrics of the ResNet model.

No	Database	Accuracy	Precision	Recall	F1 Score
1	YSYW	0.84	0.75	0.75	0.75
2	Apnea-ECG	0.84	0.83	0.87	0.85

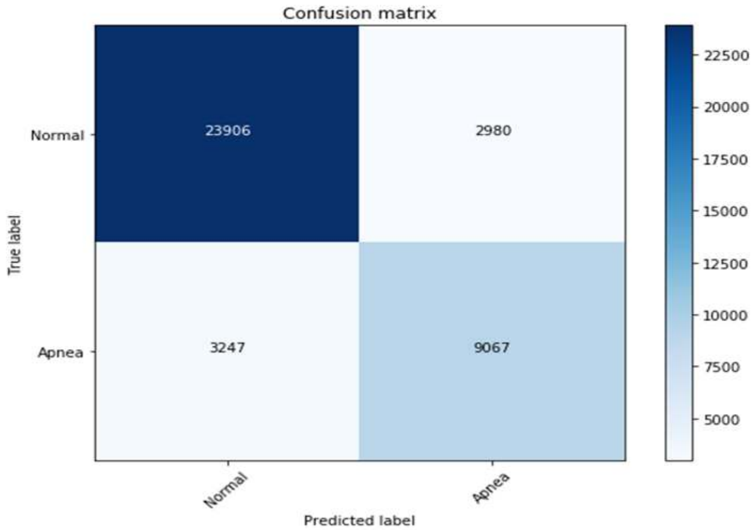

Figure 5.9. Confusion matrix of the ResNet model applied to the YSYW database. Where TP: 9067, TN: 23906, FP: 2980, FN: 3247. Normal: 0 and Apnea: 1

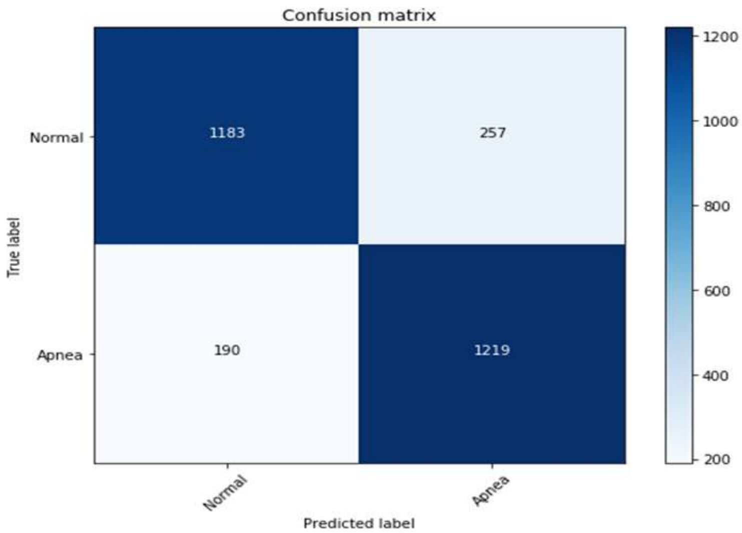

Figure 5.10. Confusion matrix of the ResNet model applied to the Apnea-ECG database. Where TP: 1219, TN: 1183, FP: 257, FN: 190. Normal: 0 and Apnea: 1

5.2.4. A Hybrid Deep Learning Model Results

We combined the previous models together in one deep learning architecture. First, one GRU layer was placed after the input layer. Second, the architecture of ResNet-50 was used. Then one bidirectional LSTM layer was placed before the output layer. We used the ReLu activation functions at the hidden layers and the sigmoid activation function at the output layer. This model gives much better results than the previous ones.

Table 5.5 consists of the evaluation metrics of the hybrid model, where accuracy is the overall accuracy (Apnea/Normal), whereas precision, recall, and F1 score belong to apnea class detection.

Table 5. 5. Evaluation metrics of the Hybrid model.

No	Database	Accuracy	Precision	Recall	F1 Score
1	YSYW	0.97	0.91	0.97	0.94
2	Apnea-ECG	0.92	0.89	0.97	0.93

Figure 5.11 and Figure 5.12 visualize the predicted output with the truth output for the YSYW and Apnea-ECG databases, respectively. Figure 5.13 and Figure 5.14 visualize an example of misclassified epochs. The confusion matrices are shown in Figure 5.15 and Figure 5.16.

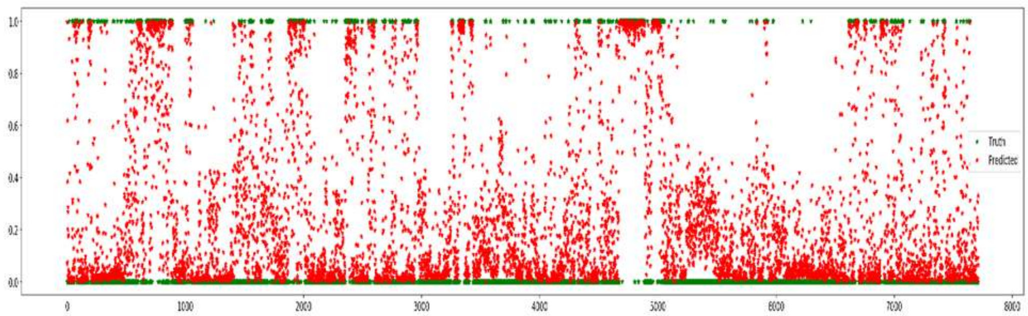

Figure 5. 11. Fitting plot for the YSYW database using the hybrid model.

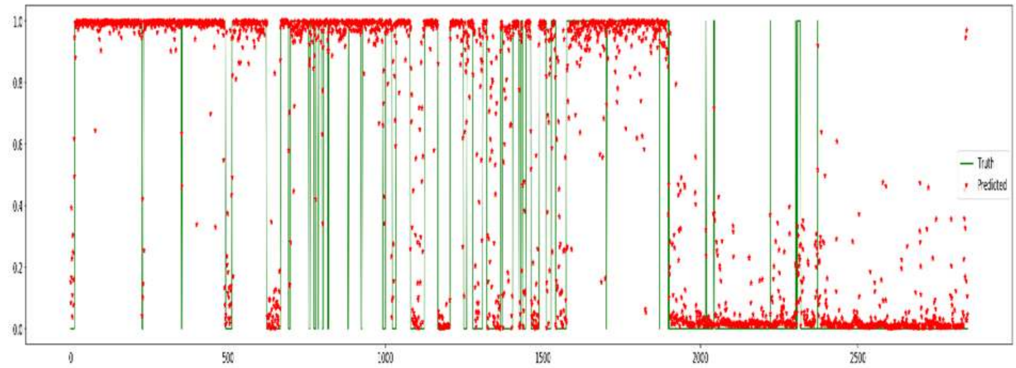

Figure 5. 12. Fitting plot for the Apnea-ECG database using the hybrid model.

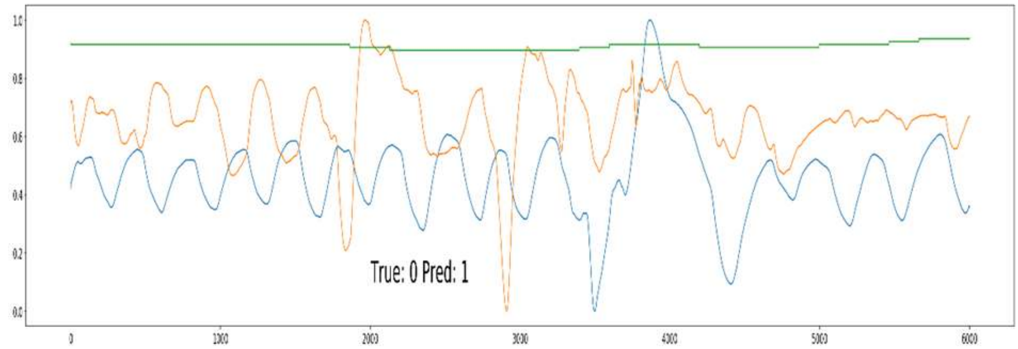

Figure 5. 13. Misclassified window (False Positive).

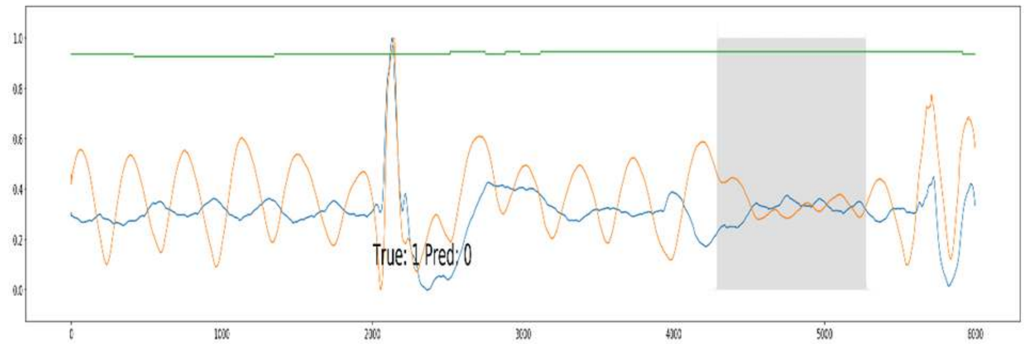

Figure 5. 14. Misclassified window (False Negative).

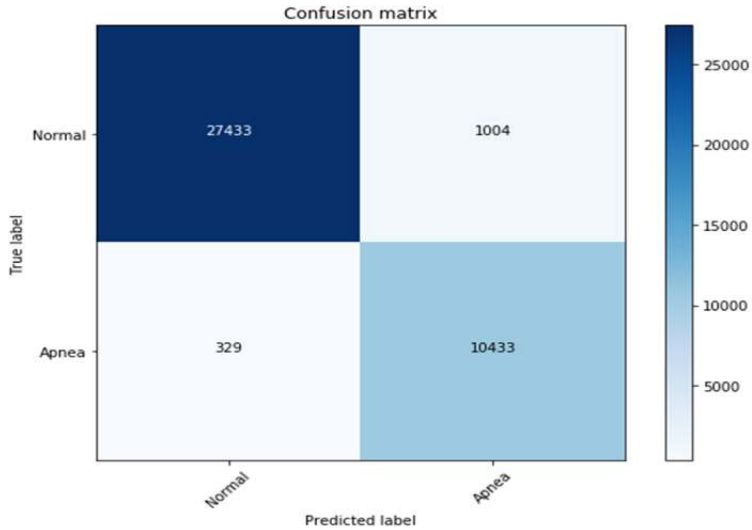

Figure 5. 15. Confusion matrix of the hybrid model applied to the YSYW database. Where TP: 10433, TN: 27433, FP: 1004, FN: 329. Normal: 0 and Apnea: 1

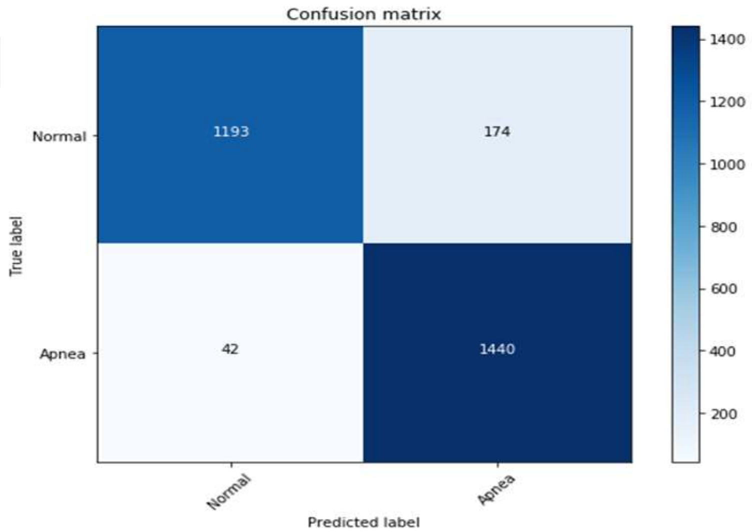

Figure 5. 16. Confusion matrix of the hybrid model applied to the Apnea-ECG database. Where TP: 1440, TN: 1193, FP: 174, FN: 42. Normal: 0 and Apnea: 1

5.3. Comparison of Methods

We can use bar and line plots to show the results graphically. A comparison between deep learning techniques utilized in the detection of sleep apnea events from PSG studies can be made according to evaluation metrics. Table 5.6. shows the results of all methods used in this thesis. The test set of the YSYW database is 91 patients because the total number of patients in the training folder is 994, and we removed the data of three patients during the data preparation due to lack of labels. Thus, the remaining set was used as 900 for the training and 91 for the test. The samples which are unlabeled were not used in this study. For the Apnea-ECG database, we mentioned in Chapter 4 that only these eight patients have related signals. We applied more than one deep learning model in order to compare the prediction and generalization performance with other methods. According to Table 5.6 and Table 5.7, it can be concluded that the hybrid model gives the best results in both databases with an overall accuracy of 97% and 92%. GRU model comes next with an overall accuracy of 92% and 91%.

Table 5. 6. Confusion matrices parameters of our models applied to the YSYW and Apnea-ECG databases.

Databases	Models	True Positive	True Negative	False Positive	False Negative
YSYW	LSTM	8507	25473	2715	2514
Apnea-ECG	LSTM	1402	1065	209	173
YSYW	GRU	9507	26622	1715	1365
Apnea-ECG	GRU	1195	1410	155	89
YSYW	ResNet50	9067	23906	2980	3247
Apnea-ECG	ResNet50	1219	1183	257	190
YSYW	Hybrid	10433	27433	1004	329
Apnea-ECG	Hybrid	1440	1193	174	42

Table 5. 7. Performance results of the deep learning models in sleep apnea events detection.

No	Databases	Signals	Models	Train	Test	Accuracy	Precision	Recall	F1-score
1	YSYW	ABD CHEST SpO2	LSTM	900	91	0.87	0.76	0.77	0.76
2	Apnea-ECG	ABD CHEST SpO2	LSTM	-	8	0.87	0.87	0.89	0.88
3	YSYW	ABD CHEST SpO2	GRU	900	91	0.92	0.85	0.87	0.86
4	Apnea-ECG	ABD CHEST SpO2	GRU	-	8	0.91	0.89	0.93	0.91
5	YSYW	ABD CHEST SpO2	ResNet	900	91	0.84	0.75	0.75	0.75
6	Apnea-ECG	ABD CHEST SpO2	ResNet	-	8	0.84	0.83	0.87	0.85
7	YSYW	ABD CHEST SpO2	Hybrid	900	91	0.97	0.91	0.97	0.94
8	Apnea-ECG	ABD CHEST SpO2	Hybrid	-	8	0.92	0.89	0.97	0.93

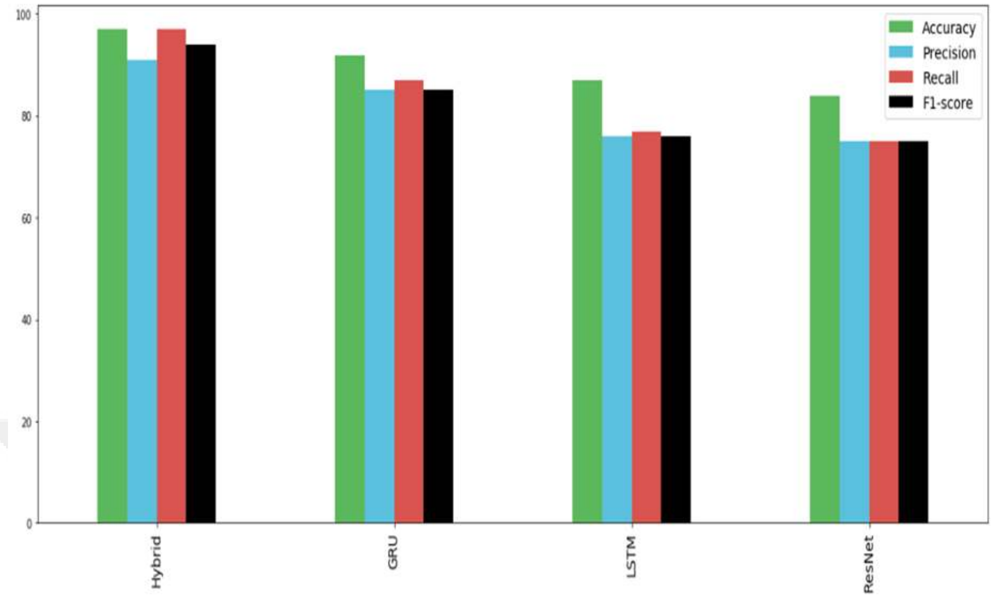

Figure 5. 17. Bar plot for results applied to the YSYW database.

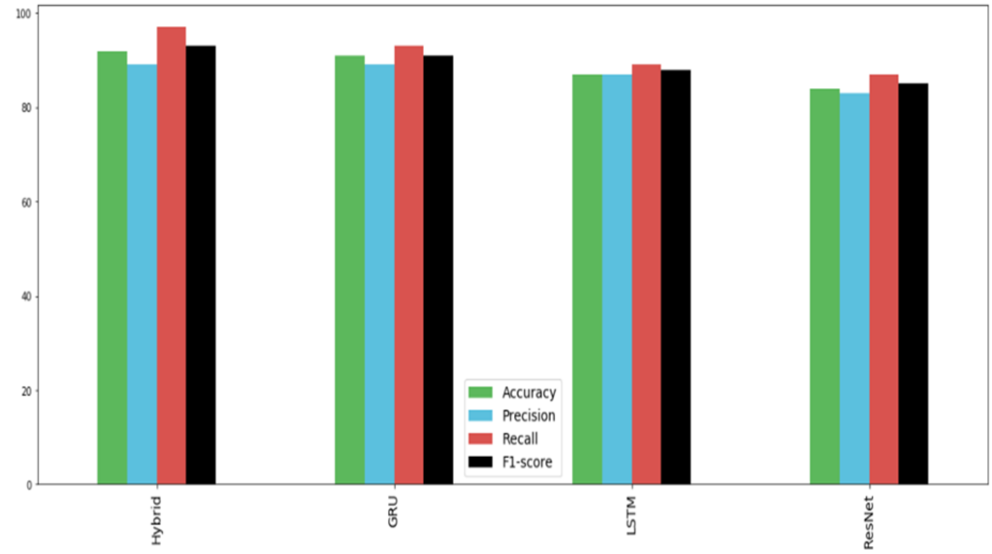

Figure 5. 18. Bar plot for results applied to the Apnea-ECG database.

For the training phase, we obtained the best validation accuracy after 700 epochs of training, as shown in Figure 5.19. By using early stopping in Keras, we controlled the validation accuracy. It was an indicator for monitoring the validation accuracy and loss. In our case, the validation accuracy started to be stable beyond 700 epochs of training. The batch size was taken as 32 because a big batch size would lead to poor generalization and lower test accuracy.

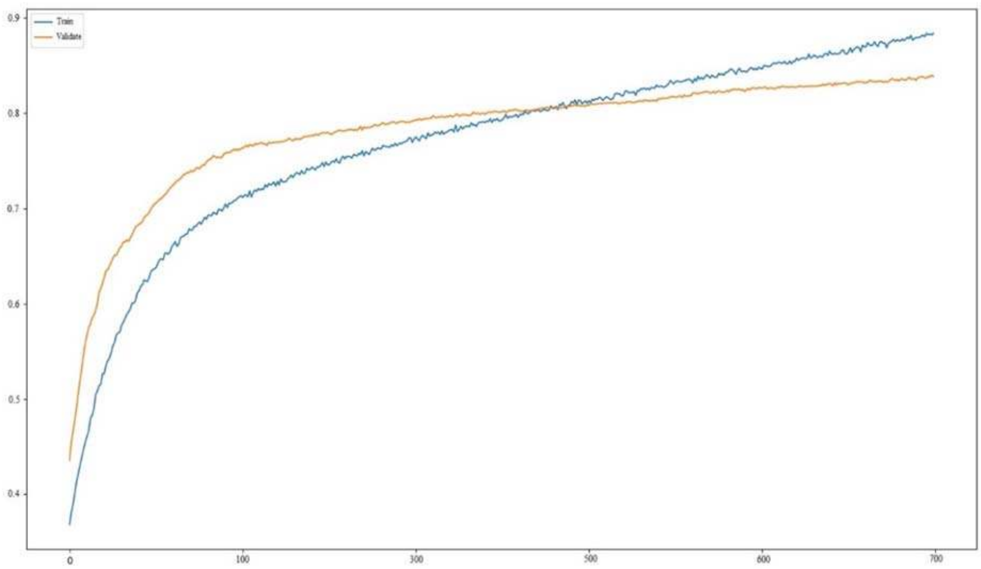

Figure 5. 19. The training and validation accuracies versus epochs.

6. DISCUSSIONS

In light of the results that we obtained in this thesis, we can say that deep learning has given high performance in this task. The four deep learning techniques, including LSTM, GRU, ResNet-50, and the hybrid model, were structured and evaluated according to several metrics.

Precision refers to exactness or quality, whereas recall refers to completeness or quantity. High precision is always an excellent indicator of successful classification for the positive class, which is the apnea class in this task.

We used the per-window method in generating signals and detection of apnea events. Some researchers used the per-record method. They computed the apnea-hypopnea index (AHI) for some windows in order to obtain a record that is classified as having normal, mild, moderate, or severe apnea.

Dropout layers played a significant role in decreasing the overfitting. The number of hidden layers or units, and the number of neurons or memory cells, were identified after several attempts. Other hyperparameters such as optimizer and activation function were selected depending on the model's final compilation.

GRU and LSTM in the hybrid model were computed unified device architecture deep neural network (CuDNN) layers, which means that GPU must be available during training. GRU has fewer parameters than LSTM. Therefore, it can be trained faster.

In practice, we found that models obtained by using only one signal from the PSG study can not be generalized for further use because the symptoms are likely to vary based on the physical variations in patients. Consequently, it was discovered that using more than one signal gives the chance to catch a higher number of abnormal events.

Nesaragi et al. (Nesaragi et al., 2018) also used the YSYW database in a different way of us. First, they focused on arousal and non-arousal events. EEG signals must be used to detect arousal events. Hypopnea and arousal events can only be distinguished with EEG signals. According to Figure 6.1, the difference between these events is shown. Second, the evaluation was not based on precision or accuracy. They opted for AUROC instead and obtained low performance. They did not discuss the internal architecture of the LSTM model. They trained two layers of quadratic discriminants (QD), which were connected to several LSTMs. Then, the output of the trained QD layers was averaged to get the final prediction.

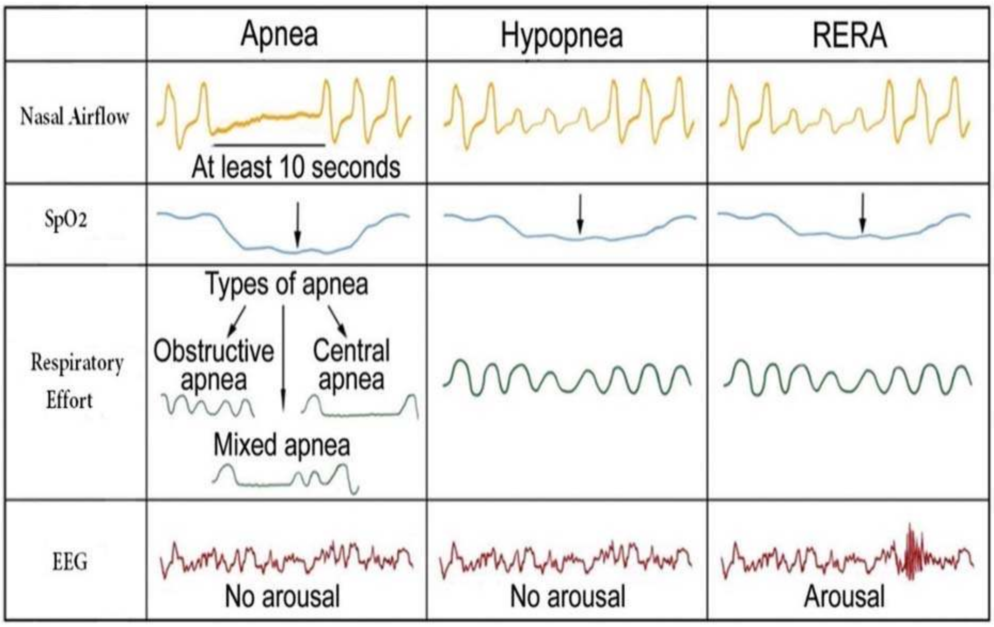

Figure 6. 1. The difference between apnea, hypopnea, and respiratory effort-related arousals (RERA).

According to Table 3.1, it can be noticed that no previous study used the YSYW database for apnea events detection. The data was published recently on Physionet, and that means our results are novel yet. Also, no previous research used the same group of signals as we used. However, most previous researchers preferred to use the Apnea-ECG database. For us, we transferred learning by loading trained models instead of training again from scratch. This approach is beneficial in many cases.

Finally, the window's size, which represents the input data, is a vital feature for increasing or decreasing the performance of the system. A window of size 30 seconds might give better results because more than one apnea might occur during one minute. That is what Pomprapa et al. (Pomprapa et al., 2018) emphasized in their paper. Figure 6.2 shows our results with those mentioned in Table 3.1.

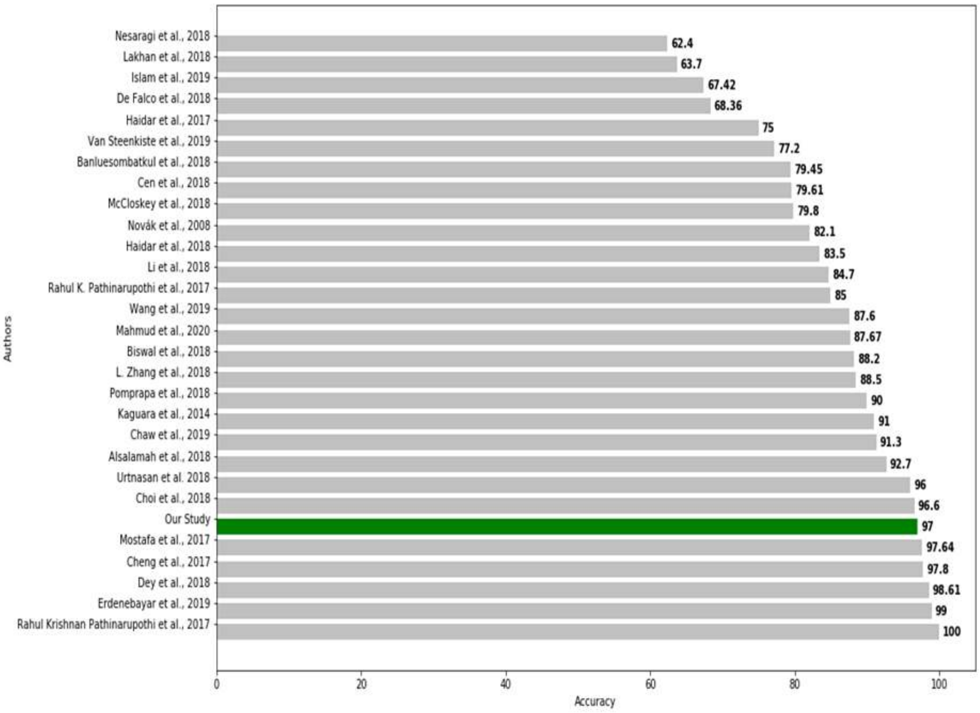

Figure 6. 2. Lollipop plot in which emphasizes our results.

7. CONCLUSIONS

Sleep apnea syndrome is a serious disease with complaints that cannot be relieved without treatment. The condition is diagnosed with difficulty because it only occurs during sleep. The existing diagnosis technique requires an excessive effort with respect to the patient and a sleep expert. The cost and time of diagnosis are exorbitant. Hence, neural network techniques can be used in this field to provides the necessary solutions.

In this thesis, we provided an overview of the cause of the problem and the motivation behind it. We thoroughly detailed the disease and its types. It was observed that the most common type to which treatment techniques are applied is OSA.

While reviewing previous researches, we noticed that deep learning became widely used in projects where the data type is physiological signals. The data taken from human bodies require a high level of experience to handle it. Sleep apnea is diagnosed by monitoring and analysis every 30 or 60 seconds of the total sleep time. Thus, deep learning saves efforts.

The trained model of this study can be reused in a sleep lab or home test. The patient can collect the same data that we used by using sensors, which can be easily obtained. On the other hand, sleep technologists can also compare their diagnosis with the predicted diagnosis to improve accuracy.

As future work, we will study to make these models in such a way that they can detect all other diseases related to sleep, not just sleep apnea. We can also detect sleep stages, which are very important for researchers and sleep technicians in this field. Sleep stages can be obtained by analyzing EEG signals, which show brain activities. EEG is the key to diagnose other familiar diseases such as epilepsy. It is considered as the only effective way that determines whether a person is asleep or awake at the time of the sleep apnea test.

In conclusion, collecting diagnosed patients' data is the most critical problem that faced researchers in this field. Literary, the machine's intelligence depends on data quality and annotation's accuracy. In some cases, more than one annotation may be used to get a record well classified.

REFERENCES

360healthcity.com. 2019. "SNORING: HOW TO STOP SNORING | EVERYTHING YOU NEED TO KNOW." Retrieved (https://360healthcity.com/snoring-everything-you-need-to-know/).

Adf.org.au. 2020. "Opium." Retrieved (https://adf.org.au/drug-facts/opium/).

Aerjournal.com. 2019. "Relationship Between Obstructive Sleep Apnoea And AF." Retrieved (https://www.aerjournal.com/articles/relationship-between-obstructive-sleep-apnoea-af).

Almazaydeh, Laiali, Khaled Elleithy, and Miad Faezipour. 2012. "Detection of Obstructive Sleep Apnea through ECG Signal Features." in IEEE International Conference on Electro Information Technology.

Alsalamah, Mashail, Saad Amin, and Vasile Palade. 2018. "Detection of Obstructive Sleep Apnea Using Deep Neural Network." Pp. 97–120 in Applications of Big Data Analytics: Trends, Issues, and Challenges. Springer International Publishing.

American Heart Association. 2017. "Understanding Blood Pressure Readings | American Heart Association." Aha 01. Retrieved (https://www.heart.org/en/health-topics/high-blood-pressure/understanding-blood-pressure-readings).

Aurora, R. Nisha, Susmita Chowdhuri, Kannan Ramar, Sabin R. Bista, Kenneth R. Casey, Carin I. Lamm, David A. Kristo, Jorge M. Mallea, James A. Rowley, Rochelle S. Zak, and Sharon L. Tracy. 2012. "The Treatment of Central Sleep Apnea Syndromes in Adults: Practice Parameters with an Evidence-Based Literature Review and Meta-Analyses." Sleep 35(1): 17–40.

Banluesombatkul, Nannapas, Thanawin Rakthanmanon, and Theerawit Wilaiprasitporn. 2018. "Single Channel ECG for Obstructive Sleep Apnea Severity Detection Using a Deep Learning Approach." IEEE Region 10 Annual International Conference, Proceedings/TENCON 2018-Octob:2011–16.

Bisong, Ekaba. 2019. "Google Colaboratory." Pp. 59–64 in Building Machine Learning and Deep Learning Models on Google Cloud Platform. Apress.

Biswal, Siddharth, Haoqi Sun, Balaji Goparaju, M. Brandon Westover, Jimeng Sun, and Matt T. Bianchi. 2018. "Expert-Level Sleep Scoring with Deep Neural Networks." Journal of the American Medical Informatics Association 25(12): 1643–50.

Blank, Janet Babich, Peggy Mannen Cawthon, Mary Lou Carrion-Petersen, Loretta Harper, J. Phillip Johnson, Eileen Mitson, and Romelia Ramírez Delay. 2005. "Overview of Recruitment for the Osteoporotic Fractures in Men Study (MrOS)." Contemporary Clinical Trials 26(5): 557–68.

Blumen, Marc B., Annie Perez De La Sota, Maria A. Quera-Salva, Bruno Frachet, Frédéric Chabolle, and Frédéric Lofaso. 2004. "Tongue Mechanical Characteristics and Genioglossus Muscle EMG in Obstructive Sleep Apnoea Patients." Respiratory Physiology and Neurobiology 140(2): 155–64.

Carissa Stephens. 2017. "How to Check Your Pulse." Medical News Today . Retrieved (https://www.medicalnewstoday.com/articles/258118).

Cen, Ling, Zhu Liang Yu, Tilmann Kluge, and Wee Ser. 2018. "Automatic System for Obstructive Sleep Apnea Events Detection Using Convolutional Neural Network." Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2018-July(Ml): 3975–78.

Chaw, Hnin Thiri, Sinchai Kamolphiwong, and Krongthong Wongsritrang. 2019. "Sleep Apnea Detection Using Deep Learning." Tehnički Glasnik 13(4): 261–66.

Cheng, Maowei, Worku J. Sori, Feng Jiang, Adil Khan, and Shaohui Liu. 2017. "Recurrent Neural Network Based Classification of ECG Signal Features for Obstruction of Sleep Apnea Detection." Proceedings - 2017 IEEE International Conference on Computational Science and Engineering and IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, CSE and EUC 2017 2: 199–202.

Cho, Kyunghyun, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. "Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation." Pp. 1724–34 in EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference. Association for Computational Linguistics (ACL).

Choi, Sang Ho, Heenam Yoon, Hyun Seok Kim, Han Byul Kim, Hyun Bin Kwon, Sung Min Oh, Yu Jin Lee, and Kwang Suk Park. 2018. "Real-Time Apnea-Hypopnea Event Detection during Sleep by Convolutional Neural Networks." Computers in Biology and Medicine 100: 123–31.

Dean, Dennis A., Ary L. Goldberger, Remo Mueller, Matthew Kim, Michael Rueschman, Daniel Mobley, Satya S. Sahoo, Catherine P. Jayapandian, Licong Cui, Michael G. Morrical, Susan Surovec, Guo-Qiang Zhang, and Susan Redline. 2016. "Scaling Up Scientific Discovery in Sleep Medicine: The National Sleep Research Resource." Sleep 39(5): 1151–64.

Dey, Debangshu, Sayanti Chaudhuri, and Sugata Munshi. 2018. "Obstructive Sleep Apnoea Detection Using Convolutional Neural Network Based Deep Learning Framework." Biomedical Engineering Letters 8(1): 95–100.

Dprogrammer.org. 2020. "DProgrammer Lopez." Retrieved (http://dprogrammer.org/rnn-lstm-gru).

Erdenebayar, Urtnasan, Yoon Ji Kim, Jong Uk Park, Eun Yeon Joo, and Kyoung Joung Lee. 2019. "Deep Learning Approaches for Automatic Detection of Sleep Apnea Events from an Electrocardiogram." Computer Methods and Programs in Biomedicine 180:105001.

De Falco, Ivanoe, Giuseppe De Pietro, Giovanna Sannino, Umberto Scafuri, Ernesto Tarantino, Antonio Della Cioppa, and Giuseppe A. Trunfio. 2018. “Deep Neural Network Hyper-Parameter Setting for Classification of Obstructive Sleep Apnea Episodes.” Proceedings - IEEE Symposium on Computers and Communications 2018-June: 1187–92.

Ghassemi, Mohammad M., Benjamin E. Moody, Li Wei H. Lehman, Christopher Song, Qiao Li, Haoqi Sun, Roger G. Mark, M. Brandon Westover, and Gari D. Clifford. 2018. “You Snooze, You Win: The PhysioNet/Computing in Cardiology Challenge 2018.” Computing in Cardiology 2018-Septe: 20–23.

Goldberger, A. L., L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. 2000. “PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals.” Circulation 101(23).

Guan, Lili, Weiliang Wu, Yating Huo, Xiaoying Li, Bingpeng Guo, Yuqiong Yang, Jiawen Xu, Fangfang Jiang, Luqian Zhou, and Rongchang Chen. 2018. “Efficacy of Bilevel Positive Airway Pressure and Continuous Positive Airway Pressure Therapy in Patients with Obesity Hypoventilation Syndrome: Protocol for Systematic Review and Meta-Analysis.” BMJ Open 8(5).

Haidar, Rim, Irena Koprinska, and Bryn Jeffries. 2017. “Sleep Apnea Event Detection from Nasal Airflow Using Convolutional Neural Networks.” Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 10638 LNCS: 819–27.

Haidar, Rim, Stephen McCloskey, Irena Koprinska, and Bryn Jeffries. 2018. “Convolutional Neural Networks on Multiple Respiratory Channels to Detect Hypopnea and Obstructive Apnea Events.” Proceedings of the International Joint Conference on Neural Networks 2018-July: 1–7.

Health.clevelandclinic.org. 2017. "Why Sleep Apnea Raises Your Risk of Sudden Cardiac Death." Retrieved (https://health.clevelandclinic.org/why-sleep-apnea-raises-your-risk-of-sudden-cardiac-death/).

Hochreiter, Sepp, and Jürgen Schmidhuber. 1997. "Long Short-Term Memory." Neural Computation 9(8): 1735–80.

HRUBOS-STRØM, HARALD, ANNA RANDBY, SILJE K. NAMTVEDT, HÅVARD A. KRISTIANSEN, GUNNAR EINVIK, JŪRATĖŠALTYTĖ BENTH, VIREND K. SOMERS, INGER H. NORDHUS, MICHAEL B. RUSSELL, TORIL DAMMEN, TORBJØRN OMLAND, and KARI J. KVAERNER. 2011. "A Norwegian Population-Based Study on the Risk and Prevalence of Obstructive Sleep Apnea The Akershus Sleep Apnea Project (ASAP)." Journal of Sleep Research 20(1pt2): 162–70.

Islam, Syed Ms, Hassan Mahmood, Adel Ali Al-Jumaily, and Scott Claxton. 2019. "Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction." Proceedings - International Conference on Machine Learning and Data Engineering, ICMLDE 2018 154–57.

Kaguara, Antony, Kee Myoung Nam, and Siddarth Reddy. 2014. "A Deep Neural Network Classifier for Diagnosing Sleep Apnea from ECG Data on Smartphones and Small Embedded Systems."

Kapur, Vishesh K., Dennis H. Auckley, Susmita Chowdhuri, David C. Kuhlmann, Reena Mehra, Kannan Ramar, and Christopher G. Harrod. 2017. "Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea." Journal of Clinical Sleep Medicine 1313(3): 479–504.

Karakis, Ioannis, Keith H. Chiappa, Marta San Luciano, Kenneth C. Sassower, John W. Stakes, and Andrew J. Cole. 2012. "The Utility of Routine EEG in the Diagnosis of Sleep Disordered Breathing." Pp. 333–38 in Journal of Clinical Neurophysiology. Vol. 29.

Khan, Muhammad Talha, and Rose Amy Franco. 2014. "Complex Sleep Apnea Syndrome." Hindawi Publishing Corporation - Sleep Disorders 2014:6.

Lakhan, Payongkit, Apiwat Ditthapron, Nannapas Banluesombatkul, and Theerawit Wilaiprasitporn. 2018. "Deep Neural Networks with Weighted Averaged Overnight Airflow Features for Sleep Apnea-Hypopnea Severity Classification." IEEE Region 10 Annual International Conference, Proceedings/TENCON 2018-Octob:441–45.

Li, Kunyang, Weifeng Pan, Yifan Li, Qing Jiang, and Guanzheng Liu. 2018. "A Method to Detect Sleep Apnea Based on Deep Neural Network and Hidden Markov Model Using Single-Lead ECG Signal." Neurocomputing 294: 94–101.

Mahmud, Tanvir, Md. Abu Aeioub Ansary, Talha Ibn Mahmud, Ishtiaque Ahmed Khan, and Shaikh Anowarul Fattah. 2020. "Real Time Sleep Apnea Event Detection with Deep Neural Network." 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON) 37–41.

mayoclinic.org. 2020. "Obstructive Sleep Apnea - Diagnosis and Treatment." Retrieved (https://www.mayoclinic.org/diseases-conditions/obstructive-sleep-apnea/diagnosis-treatment/drc-20352095).

McCloskey, Stephen, Rim Haidar, Irena Koprinska, and Bryn Jeffrie. 2018. "Detecting Hypopnea and Obstructive Apnea Events Using Convolutional Neural Networks on Wavelet Spectrograms of Nasal Airflo." Phung D., Tseng V., Webb G., Ho B., Ganji M., Rashidi L. (Eds) Advances in Knowledge Discovery and Data Mining. PAKDD 2018. Lecture Notes in Computer Science 10937(June): 361–72.

Michels, Daniel de Sousa, Amanda da Mota Silveira, Márcio Nakanishi Rodrigues, André Luiz Lopes Sampaio, and Alessandra Ramos Venosa. 2014. "Nasal Involvement in Obstructive Sleep Apnea Syndrome." International Journal of Otolaryngology 2014.

Mokhlesi, Babak, and Naresh M. Punjabi. 2012. "'REM-Related' Obstructive Sleep Apnea: An Epiphenomenon or a Clinically Important Entity?" Sleep 35(1): 5–7.

Mostafa, Sheikh Shanawaz, Fábio Mendonça, Fernando Morgado-Dias, and Antonio Ravelo-García. 2017. "SpO2 Based Sleep Apnea Detection Using Deep Learning." INES 2017 - IEEE 21st International Conference on Intelligent Engineering Systems, Proceedings 2017-Janua: 91–96.

ncbi.nlm.nih.gov. 2020. "Sleep Apnea." Retrieved (https://www.nhlbi.nih.gov/health-topics/sleep-apnea).

Nesaragi, Naimahmed, Ashish Sharma, Shivnarayan Patidar, Shubha Majumder, and Kouhyar Tavakolian. 2018. "Application of Recurrent Neural Network for the Prediction of Target Non-Apneic Arousal Regions in Physiological Signals." Computing in Cardiology 45: 2–5.

Novák, D., K. Mucha, and T. Al-Ani. 2008. "Long Short-Term Memory for Apnea Detection Based on Heart Rate Variability." Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'08 - "Personalized Healthcare through Technology" 5234–37.

Pagel, J. F., Carol Kwiatkowski, and Bennet Parnes. 2011. "The Effects of Altitude Associated Central Apnea on the Diagnosis and Treatment of Obstructive Sleep Apnea: Comparative Data from Three Different Altitude Locations in the Mountain West." Journal of Clinical Sleep Medicine 7(6): 610–15.

Pan, Sinno Jialin, and Qiang Yang. 2009. "A Survey on Transfer Learning."

Pathinarupothi, Rahul K., R. Vinaykumar, Ekanath Rangan, E. Gopalakrishnan, and K. P. Soman. 2017. "Instantaneous Heart Rate as a Robust Feature for Sleep Apnea Severity Detection Using Deep Learning." 2017 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2017 (April):293–96.

Pathinarupothi, Rahul Krishnan, J. Dhara Prathap, Ekanath Srihari Rangan, A. E. Gopalakrishnan, R. Vinaykumar, and K. P. Soman. 2017. "Single Sensor Techniques for Sleep Apnea Diagnosis Using Deep Learning." Proceedings - 2017 IEEE International Conference on Healthcare Informatics, ICHI 2017 524–29.

Peng, Jie, Shuai Kang, Zhengyuan Ning, Hangxia Deng, Jingxian Shen, Yikai Xu, Jing Zhang, Wei Zhao, Xinling Li, Wuxing Gong, Jinhua Huang, and Li Liu. 2020. "Residual Convolutional Neural Network for Predicting Response of Transarterial Chemoembolization in Hepatocellular Carcinoma from CT Imaging." European Radiology 30(1): 413–24.

Penzel, T., G. B. Moody, R. G. Mark, A. L. Goldberger, and J. H. Peter. 2000. "Apnea-ECG Database." Computers in Cardiology 255–58.

Physionet. 2008. "St. Vincent's University Hospital/University College Dublin Sleep Apnea Database." Retrieved (https://archive.physionet.org/pn3/ucddb/).

PhysioNet. 2002. "MGH/MF Waveform Database." Retrieved (https://physionet.org/content/mghdb/1.0.0/).

PhysioNet. 2005. "MIT-BIH Arrhythmia Database." Retrieved (https://physionet.org/content/mitdb/1.0.0/).

Pinto, Venessa L., and Sandeep Sharma. 2020. Continuous Positive Airway Pressure (CPAP). StatPearls Publishing.

Pomprapa, A., M. S. Sayani, T. Anwar, A. Stollenwerk, S. Kowalewski, P. Von Platen, and S. Leonhardt. 2018. "Apnea Detection in a Contactless Multisensor System Using Deep Learning Algorithm." in 13th Russian-German Conference on Biomedical Engineering (RGC). RWTH.

Quan, Stuart F., V. Howard, Conrad Iber, James P. Kiley, Javier Nieto, George T. O. Connor, David M. Rapoport, Susan Redline, Iijonathan M. Samet, and Patricia W. Wahl. 1997. "The Sleep Heart Health Study: Design, Rationale, and Methods." Sleep 20(12): 1077–85.

Sciencedaily.com. 2006. "Mayo Clinic Discovers New Type Of Sleep Apnea." Retrieved (https://www.sciencedaily.com/releases/2006/09/060901161349.htm).

Sleepapnea.org. 2020. "Central Sleep Apnea." Retrieved (https://www.sleepapnea.org/learn/sleep-apnea/central-sleep-apnea/).

Van Steenkiste, Tom, Willemijn Groenendaal, DIrk Deschrijver, and Tom Dhaene. 2019. “Automated Sleep Apnea Detection in Raw Respiratory Signals Using Long Short-Term Memory Neural Networks.” IEEE Journal of Biomedical and Health Informatics 23(6): 2354–64.

Trenchea, Mihaela, Oana Deleanu, Elena Dantes, Agripina Rascu, Oana Arghir, and Paraschiva Postolache. 2016. “Smoking, Drinking, Overweight, and Obstructive Sleep Apnea Among Patients With Sleep Breathing Disorders.” Chest 149(4): A569.

Turnbull, CD, N. Petousi, D. Sen, JR Stradling, and M. Kohler. 2017. “The Effects Of Supplemental Oxygen On Blood Pressure In Obstructive Sleep Apnoea During Cpap Withdrawal.” Cochrane Central Register of Controlled Trials 72(2018 Issue 1): A202-.

Urtnasan, Erdenebayar, Jong Uk Park, Eun Yeon Joo, and Kyoung Joung Lee. 2018. “Automated Detection of Obstructive Sleep Apnea Events from a Single-Lead Electrocardiogram Using a Convolutional Neural Network.” Journal of Medical Systems 42(6).

Urtnasan, Erdenebayar, Jong Uk Park, and Kyoung Joung Lee. 2018. “Automatic Detection of Sleep-Disordered Breathing Events Using Recurrent Neural Networks from an Electrocardiogram Signal.” Neural Computing and Applications 2.

Wang, Tao, Changhua Lu, Guohao Shen, and Feng Hong. 2019. “Sleep Apnea Detection from a Single-Lead ECG Signal with Automatic Feature-Extraction through a Modified LeNet-5 Convolutional Neural Network.” PeerJ 2019(9):1–17.

Wesleypharmacy.com. 2020. “Health Issues.” Retrieved (https://www.wesleypharmacy.com/Resource-Center/Medical-Health-Issues.aspx?Issue=Sleep Apnea&topic=Overview and Facts&tid=275).

Wikihow.com. 2020. “How to Measure Oxygen Saturation Using Pulse Oximeter.” Retrieved (https://www.wikihow.com/Measure-Oxygen-Saturation-Using-Pulse-Oximeter).

Wikipedia.com. 2020. “Obstructive Sleep Apnea.” Retrieved (https://en.wikipedia.org/wiki/Obstructive_sleep_apnea).

Woehrle, Holger, Martin R. Cowie, Christine Eulenburg, Anna Suling, Christiane Angermann, Marie Pia d’Ortho, Erland Erdmann, Patrick Levy, Anita K. Simonds, Virend K. Somers, Faiez Zannad, Helmut Teschler, and Karl Wegscheider. 2017. “Adaptive Servo Ventilation for Central Sleep Apnoea in Heart Failure: SERVE-HF on-Treatment Analysis.” The European Respiratory Journal 50(2): 1601692.

Xie, Chen, and Julien Dubiel. 2016. “Native Python WFDB Package.” Retrieved (https://wfdb.readthedocs.io/en/latest/).

Zhang, Guo-Qiang, Licong Cui, Remo Mueller, Shiqiang Tao, Matthew Kim, Michael Rueschman, Sara Mariani, Daniel Mobley, and Susan Redline. 2018. “The National Sleep Research Resource: Towards a Sleep Data Commons.” Journal Ofthe American Medical Informatics Association 1351–1358.

Zhang, L., D. Fabbri, R. Upender, and D. Kent. 2018. “Automated Apnea and Hypopnea Event Detection Using Deep Learning.” Sleep 41(suppl_1): A119–A120.

BIOGRAPHY

Mahmood Najeeb ABED was born in Al-Anbar, Iraq, in 1992. He received his B.Sc. degree in Electronics and Communications Engineering Department from Al-Nahrain University in 2015. He has been working as a Data Analyst since 2018. His research areas are Neural Networks, IoT, Big Data, Digital Signal Processing, and Computer Vision.